

Lecture 03

Expressiveness & Effectiveness Basic Visualisation Techniques I

Uta Hinrichs

CS5044 – Information Visualization



University of
St Andrews

practical 1 – frequently asked questions

1. Do visualization and legend have to be in the postcard format and style as the provided visualizations?
 - You can do this as you wish. You can follow the postcard format and style, or provide a simple sketch with the legend next to it.
2. Am I required to back-up my argument with lots of additional readings on perceptual theory, etc...
 - No, a brief and concise critique according to the lecture material and book chapter is sufficient.
3. Should I comment on additional aspects in the visualization (e.g., intended themes, applied visual metaphors, or visible trends)?
 - Yes, if anything sticks out for you, you can comment on this. But you still have to stick to the word limit.

frequently asked questions

4. The visualization provided to me is already very good, and I don't know how to critique it.
 - Focus on the aspects that make it a good visualization (e.g., use of visual variables, expressiveness, effectiveness, spatial arrangement, etc...).
 - In the re-design you could try to highlight other aspects that might be of interest in the data.
5. The visualization provided to me is quite bad, and I cannot read it.
 - Describe the aspects that make it so difficult to read (e.g., use of visual variables, expressiveness, effectiveness, spatial arrangement, etc...).
 - Is there really not a single aspect that works? Is there really nothing you can decipher?
 - In the re-design you can then address these aspects that don't work.

frequently asked questions

6. Does my visualization sketch have to strictly focus on refining and improving the provided visualization, or can I come up with a sketch that focuses on other aspects (e.g., different attributes, etc...)?
 - You can provide a sketch that highlight out other aspects, but it should (1) still share at least some of the provided data attributes, and (2) you should justify your decisions carefully.
 - This type of approach makes most sense, if the provided visualization is quite good already, or if you are collecting your own data and the provided categories do not make much sense.
7. For our critique, should we comment on the design of the visualization in the context of the postcard? Some elements may not work on a small postcard but may actually work at a larger scale or with a computationally generated vis.
 - Generally, yes, the visualization should work in the context of the postcard. However, you can make a brief comment that some of the drawbacks you may encounter can be mitigated using a larger display or another presentation medium...
 - For your own visualization design you should consider a static medium (e.g., paper).

outline for today

- Expressiveness & effectiveness principles
- Basic visualisation techniques for encoding tabular data

by the end of the lecture

- Have a solid understanding of the expressiveness and effectiveness principle and be able to critique a visualisation based on this
- Have a solid understanding of
 - Scatterplots
 - Bar charts
 - Line graphs
- and what types of attributes they can represent

expressiveness & effectiveness

Channels: Expressiveness Types and Effectiveness Ranks

➔ **Magnitude Channels: Ordered Attributes**

Position on common scale



Position on unaligned scale



Length (1D size)



Tilt/angle



Area (2D size)



Depth (3D position)



Color luminance



Color saturation



Curvature



Volume (3D size)



Same

Most Effectiveness Least

➔ **Identity Channels: Categorical Attributes**

Spatial region



Color hue



Motion



Shape



choice of marks and variables

- Your choice of marks and visual variables to encode data attributes should follow 2 principles

→ Expressiveness

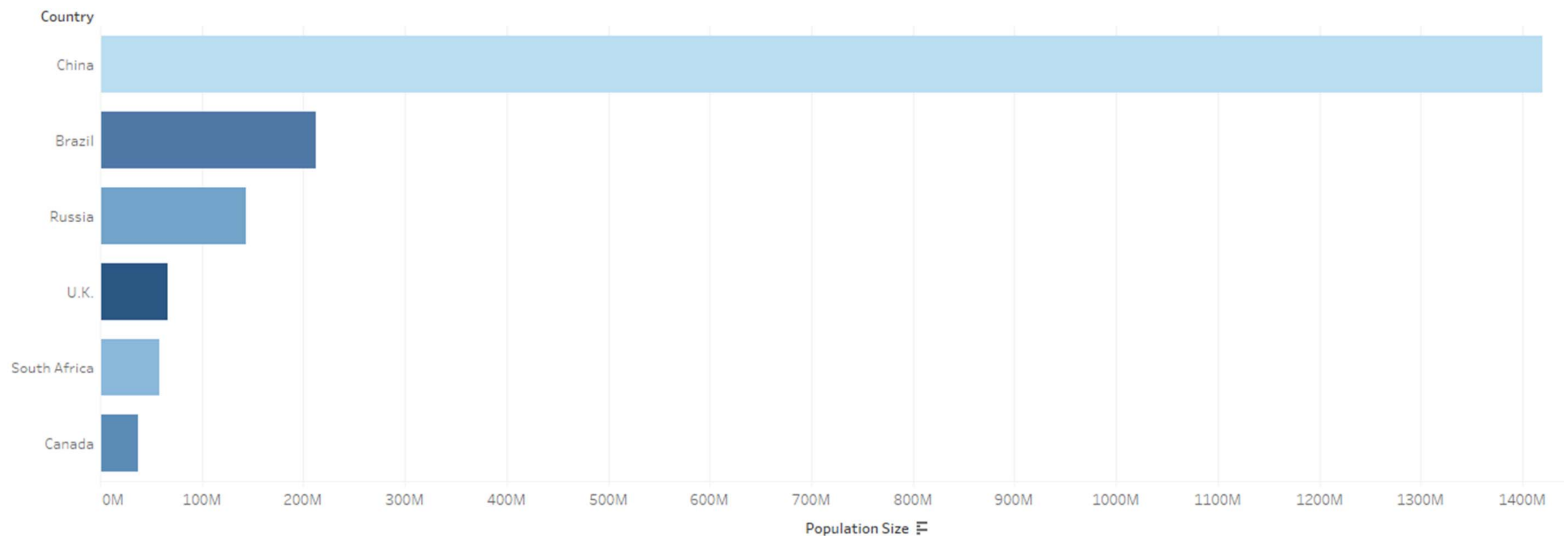
→ Efficiency

expressiveness principle

- The visual encoding should express all of, and only, the information in the dataset attributes.
- Show **ordered data** in a way that our perceptual system intrinsically senses as ordered
- Show **categorical data** in a way that does not perceptually suggest ordering.

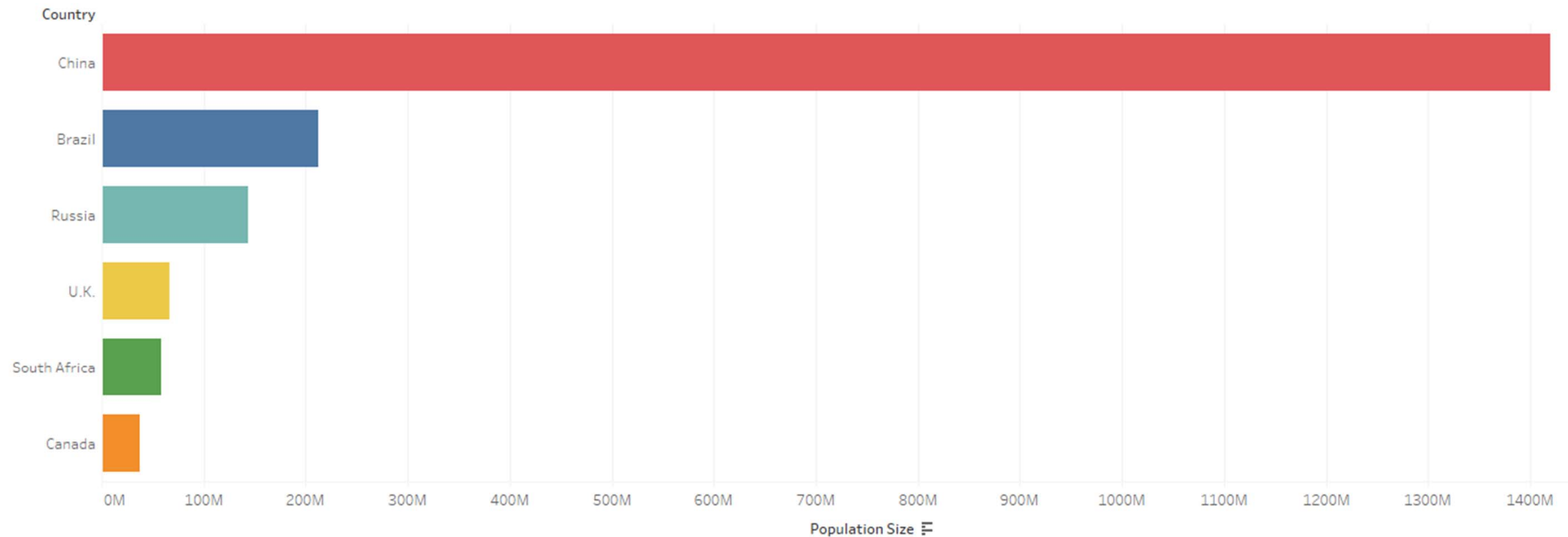
expressive if all we have is population size
and country name?

- No!
- The choice of colours implies an ordering that is not apparent in the data



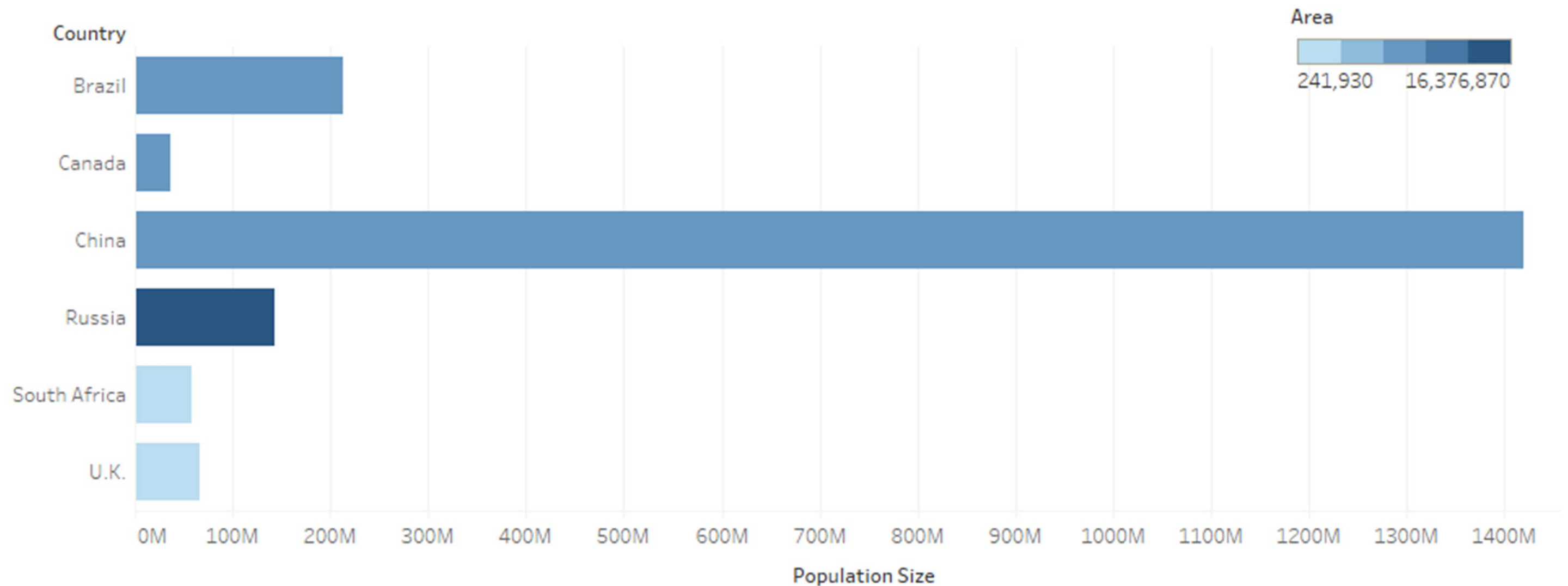
expressive if all we have is population size
and country name?

- Hue is the better choice here.



expressive: with population size, country, and area (in km²)

countryArea



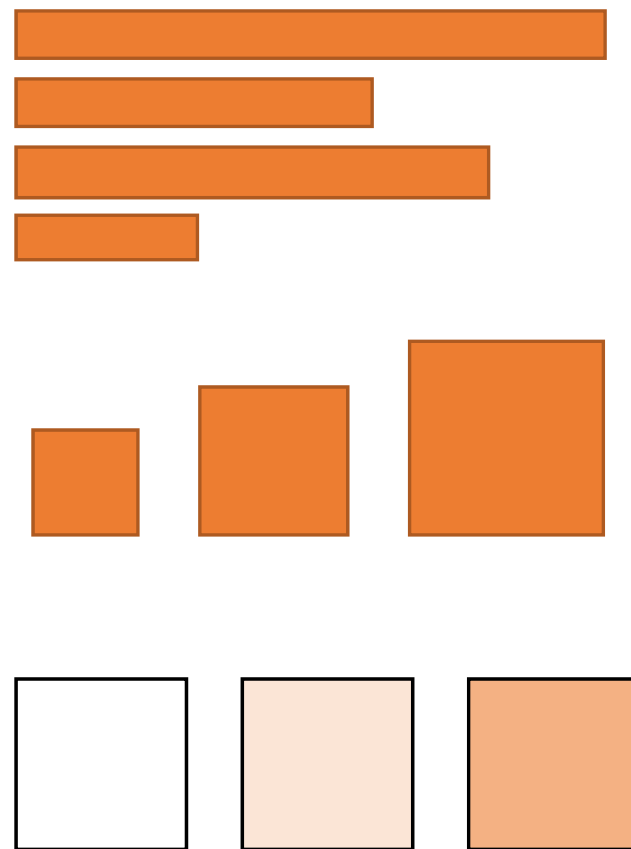
effectiveness principle

- The importance of the data attribute should be encoded **with the most effective visual variable**, in order to be noticeable.
 - Accuracy
 - Discriminability
 - Separability

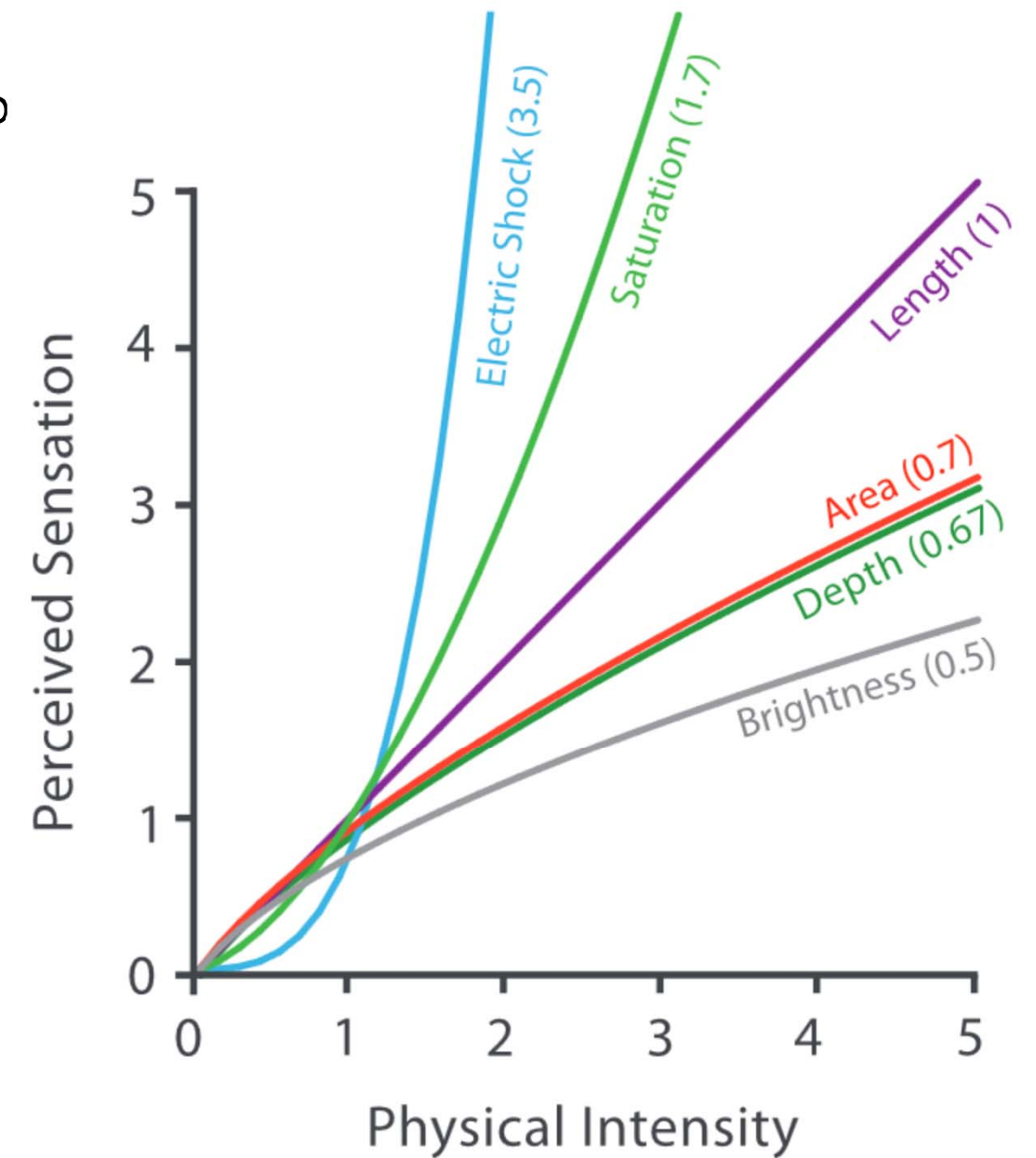
accuracy

- How accurate is our perception of a visual variable?

- Perceived sensation (S)
- Physical intensity (I)

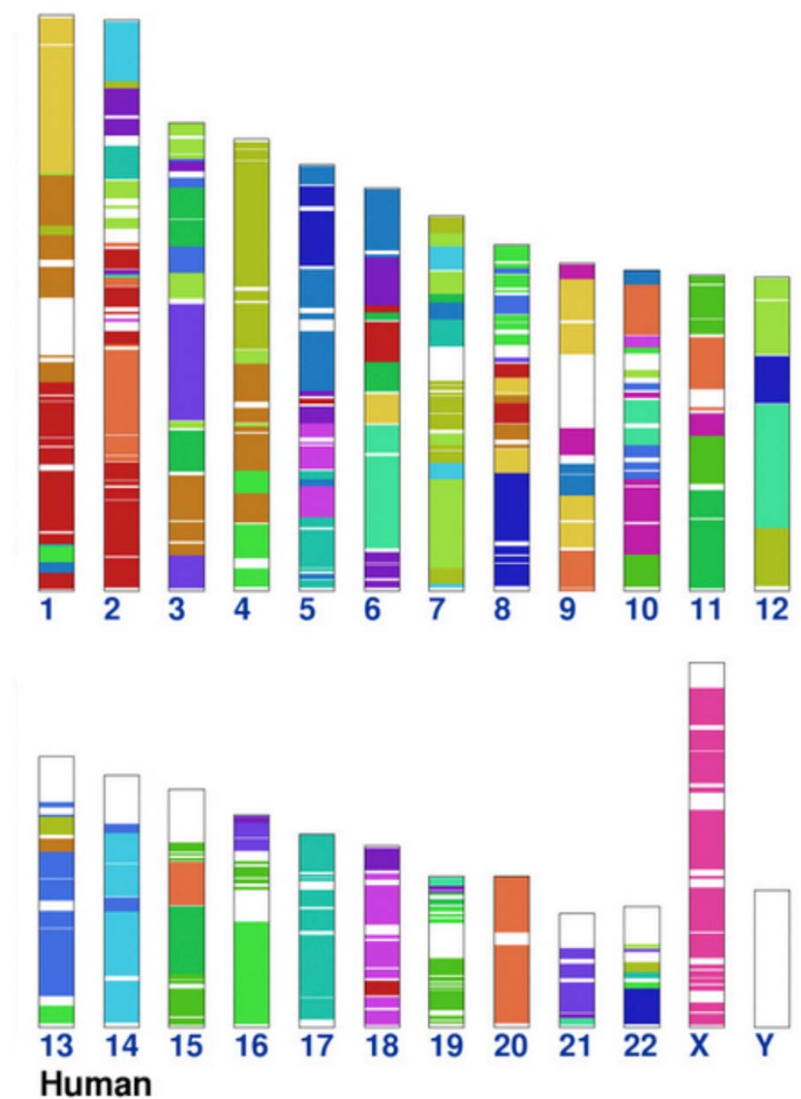


Steven's Psychophysical Power Law: $S = I^N$



discriminability

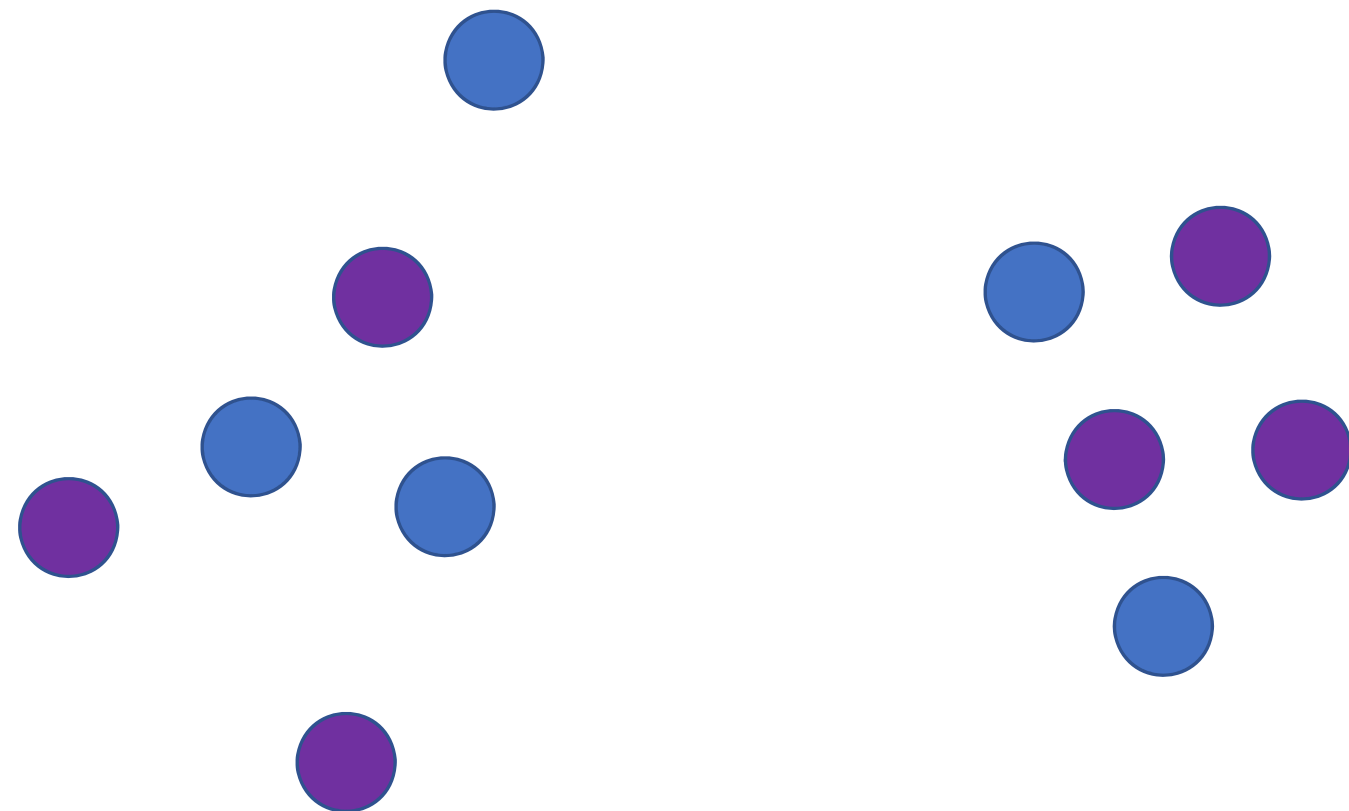
- How many distinctions are possible?
 - Colour (hue) is highly limited (7 distinctions to be sure)
 - Position is much more versatile
- Number of distinctions can depend on context



Sinha and Meller, 2007

separability

- How does the visual variable interact with other visual variables?
- Interferences between different visual variables
- Example: position + color (hue)



separability

- How does the visual variable interact with other visual variables?
- Interferences between different visual variables
- Example: size + color (hue)



pop out & grouping

→ Selective

- How easy can we identify items that stand out?
- Strong variables: position, colour, shape...
- Often depends on the number of distractors

→ Associative

- How can we see if items are similar; belong to the same group
- Strongest variable here: position

things to consider

- Carefully consider expressiveness & effectiveness of your visual representation
 - If you violate these principles, you should be able to justify your decision!
- What do you want to achieve with your visualization?
 - Accurate comparisons vs. broad explorations
- Interactivity can help to ease limitations of visual variables
 - But try not to rely on it!

expressiveness & effectiveness

visualization techniques

A PERIODIC TABLE OF VISUALIZATION METHODS

<div>>☀<</div> <div>☉</div> <div>continuum</div>	<div><div><div>>☀<</div><div>☉</div><div>Data Visualization</div><div>Visual representations of quantitative data in schematic form (either with or without axes)</div></div><div><div>>☀<</div><div>☉</div><div>Information Visualization</div><div>The use of interactive visual representations of data to amplify cognition. This means that the data is transformed into an image, it is mapped to screen space. The image can be changed by users as they proceed working with it</div></div><div><div>>☀<</div><div>☉</div><div>Concept Visualization</div><div>Methods to elaborate (mostly) qualitative concepts, ideas, plans, and analyses.</div></div></div> <div><div><div>>☀<</div><div>☉</div><div>Strategy Visualization</div><div>The systematic use of complementary visual representations in the analysis, development, formulation, communication, and implementation of strategies in organizations.</div></div><div><div>>☀<</div><div>☉</div><div>Metaphor Visualization</div><div>Visual Metaphors position information graphically to organize and structure information. They also convey an insight about the represented information through the key characteristics of the metaphor that is employed</div></div><div><div>>☀<</div><div>☉</div><div>Compound Visualization</div><div>The complementary use of different graphic representation formats in one single schema or frame</div></div></div>												<div>>☀<</div> <div>☉</div> <div>graphic facilitation</div>					
<div>>☀<</div> <div>Tb</div> <div>table</div>	<div>>☀<</div> <div>Ca</div> <div>cartesian coordinates</div>	<div>>☀<</div> <div>Me</div> <div>meeting trace</div>												<div>>☀<</div> <div>Mm</div> <div>metro map</div>	<div>☀</div> <div>Tm</div> <div>temple</div>	<div><☀></div> <div>St</div> <div>story template</div>	<div>>☀<</div> <div>Tr</div> <div>tree</div>	<div>☐</div> <div>Ct</div> <div>cartoon</div>
<div>>☀<</div> <div>Pi</div> <div>pie chart</div>	<div>>☀<</div> <div>L</div> <div>line chart</div>	<div>>☀<</div> <div>Co</div> <div>communication diagram</div>												<div>>☀<</div> <div>Fp</div> <div>flight plan</div>	<div>>☀<</div> <div>Cs</div> <div>concept sceleton</div>	<div>☀</div> <div>Br</div> <div>bridge</div>	<div>>☀<</div> <div>Fu</div> <div>funnel</div>	<div>☀</div> <div>Ri</div> <div>rich picture</div>
<div>>☀<</div> <div>B</div> <div>bar chart</div>	<div>>☀<</div> <div>Ac</div> <div>area chart</div>	<div>>☀<</div> <div>R</div> <div>radar chart cobweb</div>	<div>>☀<</div> <div>Pa</div> <div>parallel coordinates</div>	<div>>☀<</div> <div>Hy</div> <div>hyperbolic tree</div>	<div>>☀<</div> <div>Cy</div> <div>cycle diagram</div>	<div>>☀<</div> <div>T</div> <div>timeline</div>	<div>>☀<</div> <div>Ve</div> <div>venn diagram</div>	<div>>☀<</div> <div>Mi</div> <div>mindmap</div>	<div>>☀<</div> <div>Sq</div> <div>square of oppositions</div>	<div>>☀<</div> <div>Cc</div> <div>concentric circles</div>	<div>>☀<</div> <div>Ar</div> <div>argument slide</div>	<div>>☀<</div> <div>Sw</div> <div>swim lane diagram</div>	<div>>☀<</div> <div>Gc</div> <div>gantt chart</div>	<div><☉></div> <div>Pm</div> <div>perspectives diagram</div>	<div>>☉<</div> <div>D</div> <div>dilemma diagram</div>	<div><☀></div> <div>Pr</div> <div>parameter ruler</div>	<div>☀</div> <div>Kn</div> <div>knowledge map</div>	
<div>>☀<</div> <div>Hi</div> <div>histogram</div>	<div>>☀<</div> <div>Sc</div> <div>scatterplot</div>	<div>>☀<</div> <div>Sa</div> <div>sankey diagram</div>	<div>>☉<</div> <div>In</div> <div>information lense</div>	<div>>☐<</div> <div>E</div> <div>entity relationship diagram</div>	<div>>☀<</div> <div>Pt</div> <div>petri net</div>	<div>>☉<</div> <div>Fi</div> <div>flow chart</div>	<div><☀></div> <div>Cl</div> <div>clustering</div>	<div>>☀<</div> <div>Lc</div> <div>layer chart</div>	<div>>☉<</div> <div>Py</div> <div>minto pyramid technique</div>	<div>>☀<</div> <div>Ce</div> <div>cause-effect chains</div>	<div>>☀<</div> <div>Tl</div> <div>toulmin map</div>	<div>>☉<</div> <div>Dt</div> <div>decision tree</div>	<div>>☐<</div> <div>Cp</div> <div>cpm critical path method</div>	<div><☀></div> <div>Cf</div> <div>concept fan</div>	<div>>☉<</div> <div>Co</div> <div>concept map</div>	<div>☀</div> <div>Ic</div> <div>iceberg</div>	<div>☀</div> <div>Lm</div> <div>learning map</div>	
<div>>☀<</div> <div>Tk</div> <div>tukey box plot</div>	<div>>☀<</div> <div>Sp</div> <div>spectrogram</div>	<div>>☀<</div> <div>Da</div> <div>data map</div>	<div>>☉<</div> <div>Tp</div> <div>treemap</div>	<div>>☉<</div> <div>Cn</div> <div>cone tree</div>	<div>>☀<</div> <div>Sy</div> <div>system dyn./simulation</div>	<div>>☉<</div> <div>Df</div> <div>data flow diagram</div>	<div><☀></div> <div>Se</div> <div>semantic network</div>	<div>>☉<</div> <div>So</div> <div>soft system modeling</div>	<div>>☉<</div> <div>Sn</div> <div>synergy map</div>	<div><☀></div> <div>Fo</div> <div>force field diagram</div>	<div>>☐<</div> <div>Ib</div> <div>ibis argumentation map</div>	<div>>☀<</div> <div>Pr</div> <div>process event chains</div>	<div>>☀<</div> <div>Pe</div> <div>pert chart</div>	<div><☉></div> <div>Ev</div> <div>evocative knowledge map</div>	<div>>☉<</div> <div>V</div> <div>Yee diagram</div>	<div><☀></div> <div>Hh</div> <div>heaven 'n' hell chart</div>	<div>>☉<</div> <div>I</div> <div>informul</div>	

Cy Process Visualization

Hy Structure Visualization

- ☀ Overview
- ☐ Detail
- ☉ Detail AND Overview
- < > Divergent thinking
- > < Convergent thinking

<div>>☀<</div> <div>Su</div> <div>supply demand curve</div>	<div>>☀<</div> <div>Pc</div> <div>performance charting</div>	<div>>☀<</div> <div>St</div> <div>strategy map</div>	<div>>☀<</div> <div>Oc</div> <div>organisation chart</div>	<div><☐></div> <div>Ho</div> <div>house of quality</div>	<div>>☀<</div> <div>Fd</div> <div>feedback diagram</div>	<div>☐</div> <div>Ft</div> <div>failure tree</div>	<div>>☀<</div> <div>Mq</div> <div>magic quadrant</div>	<div>>☀<</div> <div>Ld</div> <div>life-cycle diagram</div>	<div>>☀<</div> <div>Po</div> <div>porter's five forces</div>	<div><☐></div> <div>S</div> <div>s-cycle</div>	<div>>☀<</div> <div>Sm</div> <div>stakeholder map</div>	<div>☉</div> <div>Is</div> <div>ishikawa diagram</div>	<div>☀</div> <div>Tc</div> <div>technology roadmap</div>
<div>☀</div> <div>Ed</div> <div>edgeworth box</div>	<div>>☀<</div> <div>Pf</div> <div>portfolio diagram</div>	<div>☀</div> <div>Sg</div> <div>strategic game board</div>	<div>>☀<</div> <div>Mz</div> <div>mintzberg's organigraph</div>	<div><☐></div> <div>Z</div> <div>zwicky's morphological box</div>	<div>>☀<</div> <div>Ad</div> <div>affinity diagram</div>	<div>☐</div> <div>De</div> <div>decision discovery diagram</div>	<div>>☀<</div> <div>Bm</div> <div>bcg matrix</div>	<div>>☀<</div> <div>Stc</div> <div>strategy canvas</div>	<div>>☀<</div> <div>Vc</div> <div>value chain</div>	<div><☐></div> <div>Hy</div> <div>hype-cycle</div>	<div>>☀<</div> <div>Sr</div> <div>stakeholder rating map</div>	<div>>☀<</div> <div>Ta</div> <div>taps</div>	<div><☐></div> <div>Sd</div> <div>spray diagram</div>

Note: Depending on your location and connection speed it can take some time to load a pop-up picture.

© Ralph Lengler & Martin J. Eppler, www.visual-literacy.org

version 1.5

http://www.visual-literacy.org/periodic_table/periodic_table.html#

⌞⌞

classifying visualization techniques

- Based on the number of attributes
- Based on the spatial arrangement

classifying vis techniques based on # of attributes

- How many attributes/dimensions/variables does the data set have?

classifying vis techniques based on # of attributes

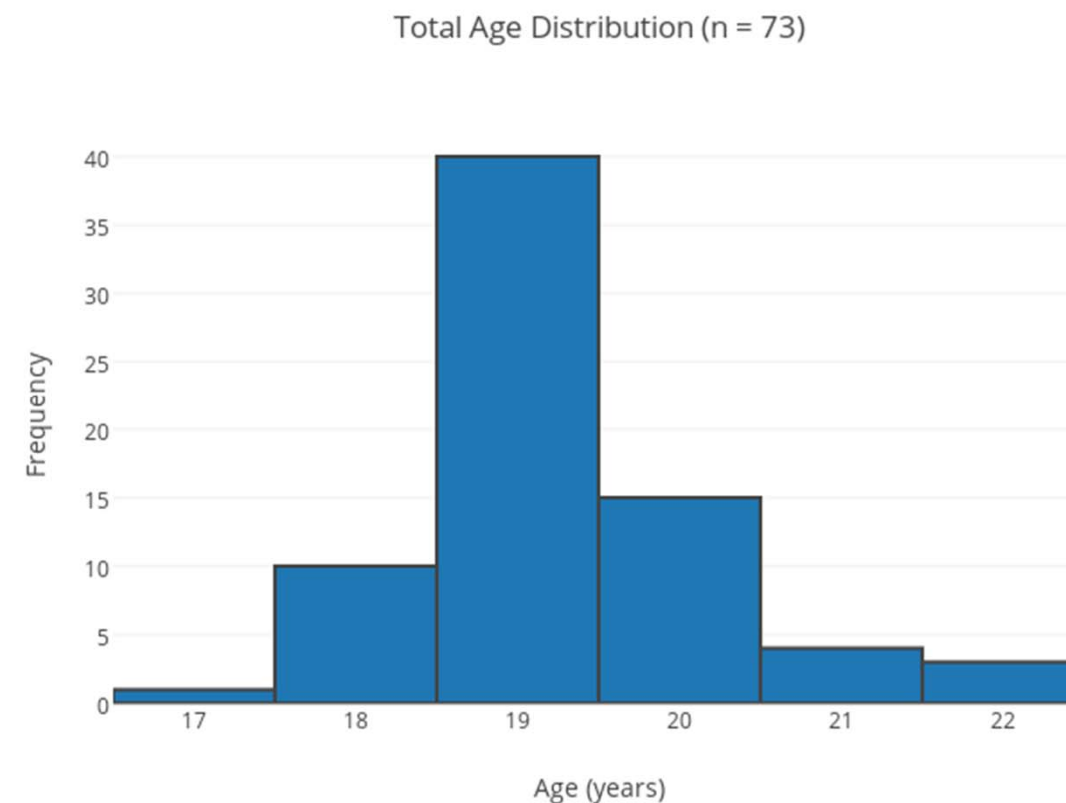
- Univariate data
 - Refer to just a single number (e.g., temperature, speed, time)



classifying vis techniques based on # of attributes

- Univariate data
 - Refer to just a single number or attribute (e.g., temperature, speed, time)
 - We may have many data along a single dimension

Id	Age
1	8
2	7
3	9
4	13
5	10



classifying vis techniques based on # of attributes

- Bivariate data
 - Data along two dimensions/variables

Id	Height	Age
1	1.10	8
2	1.00	7
3	1.20	9
4	1.70	13
5	1.20	10

classifying vis techniques based on # of attributes

- Multivariate data
 - Data along multiple dimensions

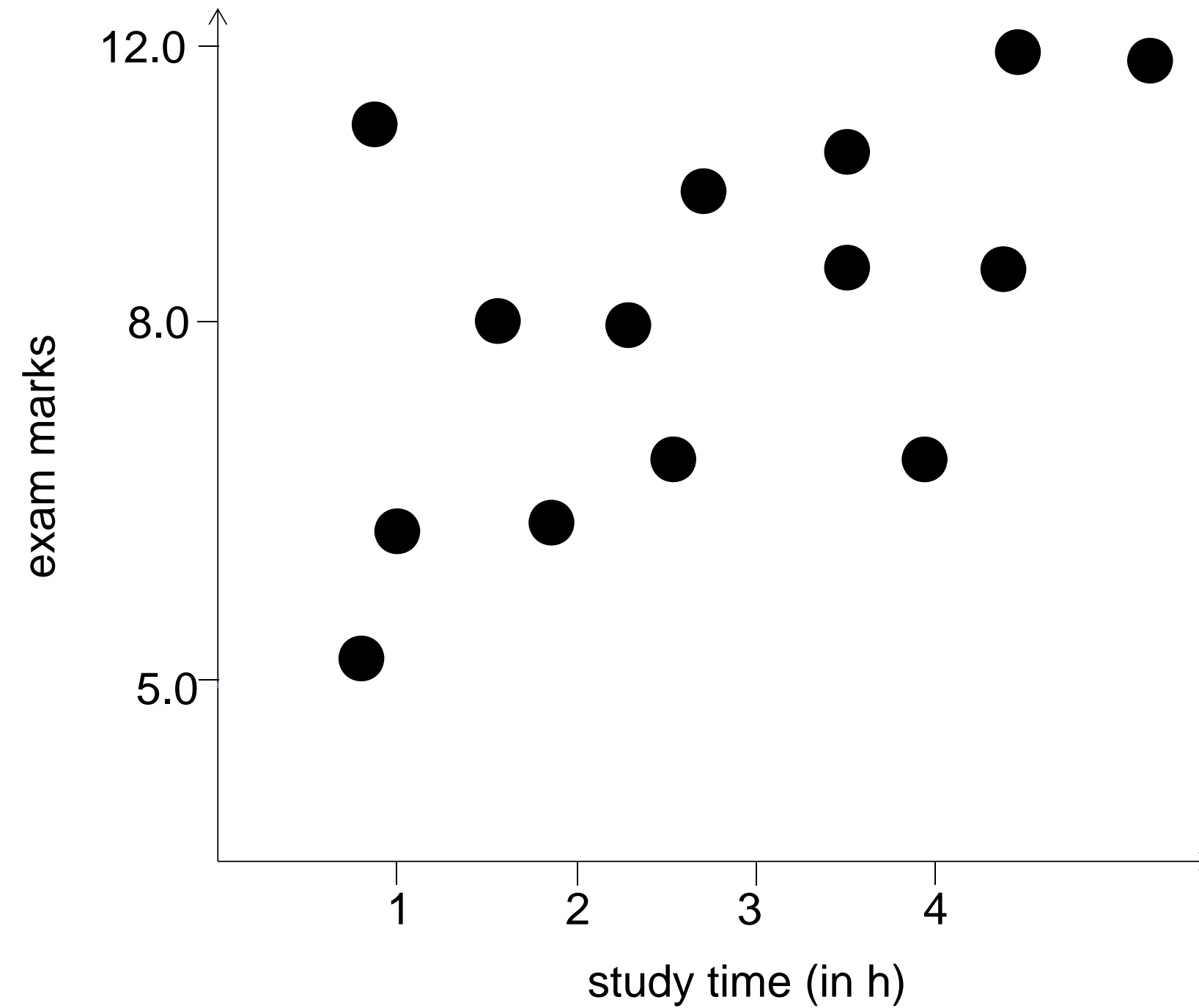
Id	Name	Height	Age	Shirt Size	Fav. Fruit
1	Apple	1.10	8	S	Banana
2	Basil	1.00	7	S	Pear
3	Clara	1.20	9	M	Durian
4	Desmond	1.70	13	L	Elderberry
5	Fanny	1.20	10	S	Lychee

classifying vis techniques based on spatial arrangement

- T. Munzner: How to arrange tabular data spatially?
 - How many dimensions do I want to show?
 - What types of attributes am I dealing with?
 - How many values do I want to show?
- Choice of visualization technique

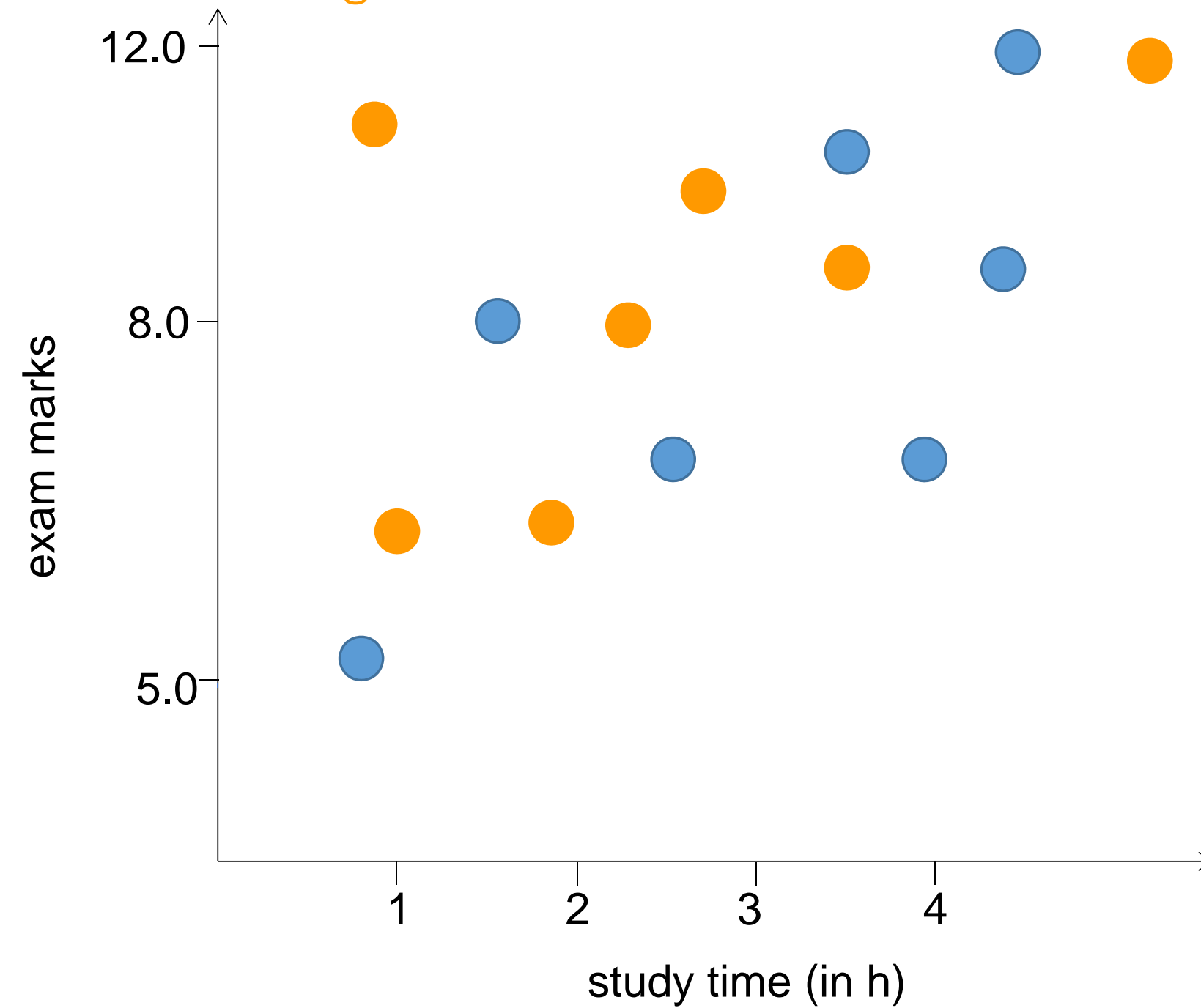
scatterplot

two quantitative attributes



scatterplot

two quantitative attributes + one categorical attribute



scatterplot / bubble chart

three quantitative attributes + one categorical attribute

- Infant mortality
- Income per person
- Population size
- Geographic region



Gapminder, income per person vs. infant mortality

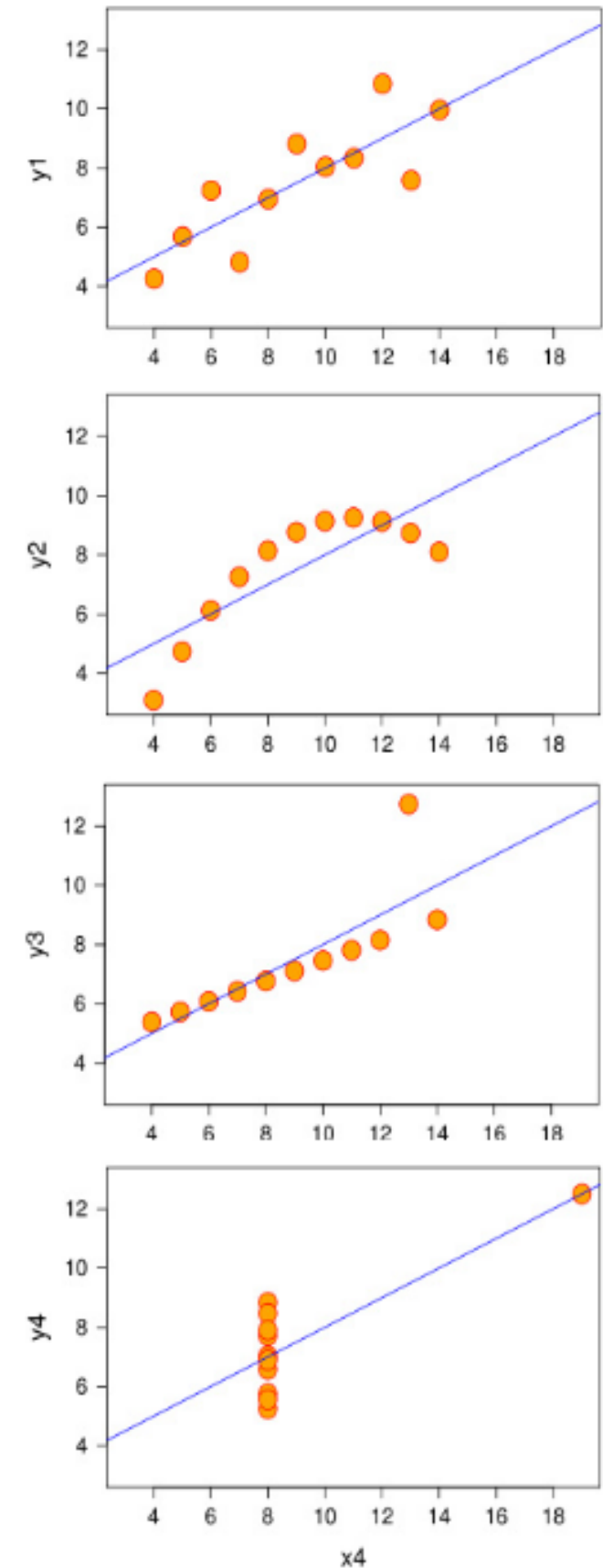
scatterplot

what Two **quantitative attributes** (at least)

how Express quantitative values as point marks using **horizontal and vertical position**

why Providing overviews

- Characterizing distributions
 - Outliers; extremes
- Finding correlations between two attributes



scatterplot

what Two **quantitative attributes** (at least)

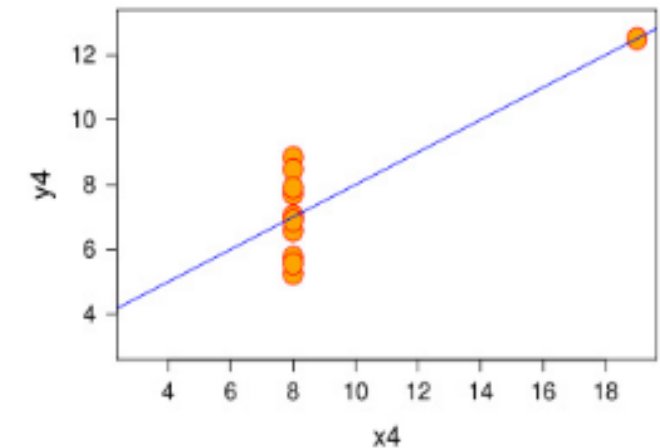
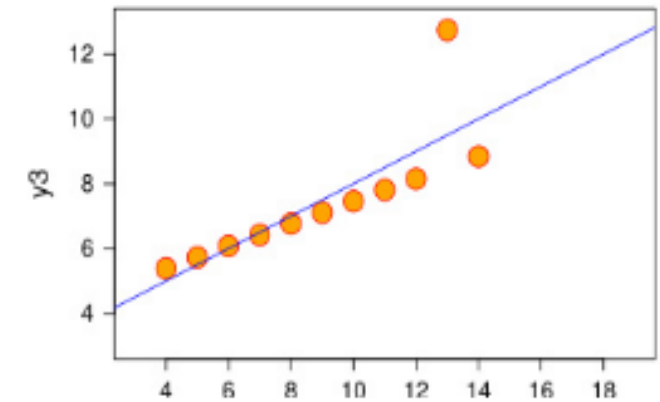
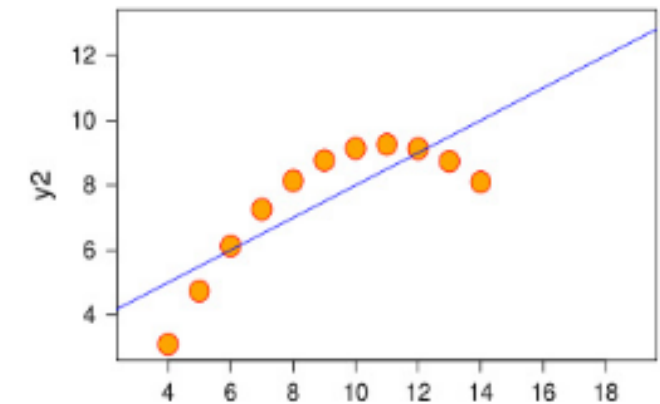
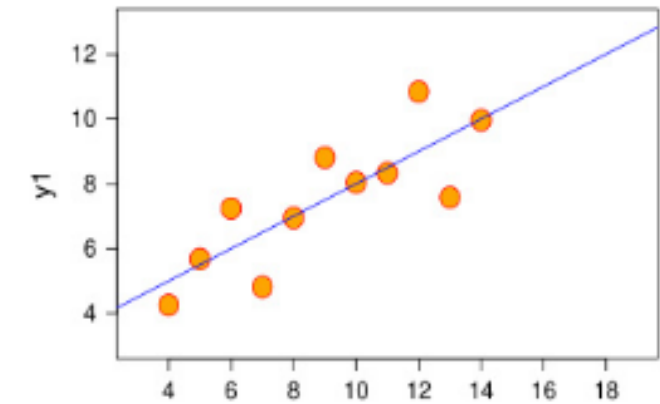
how Express quantitative values as point marks using **horizontal and vertical position**

why Providing overviews

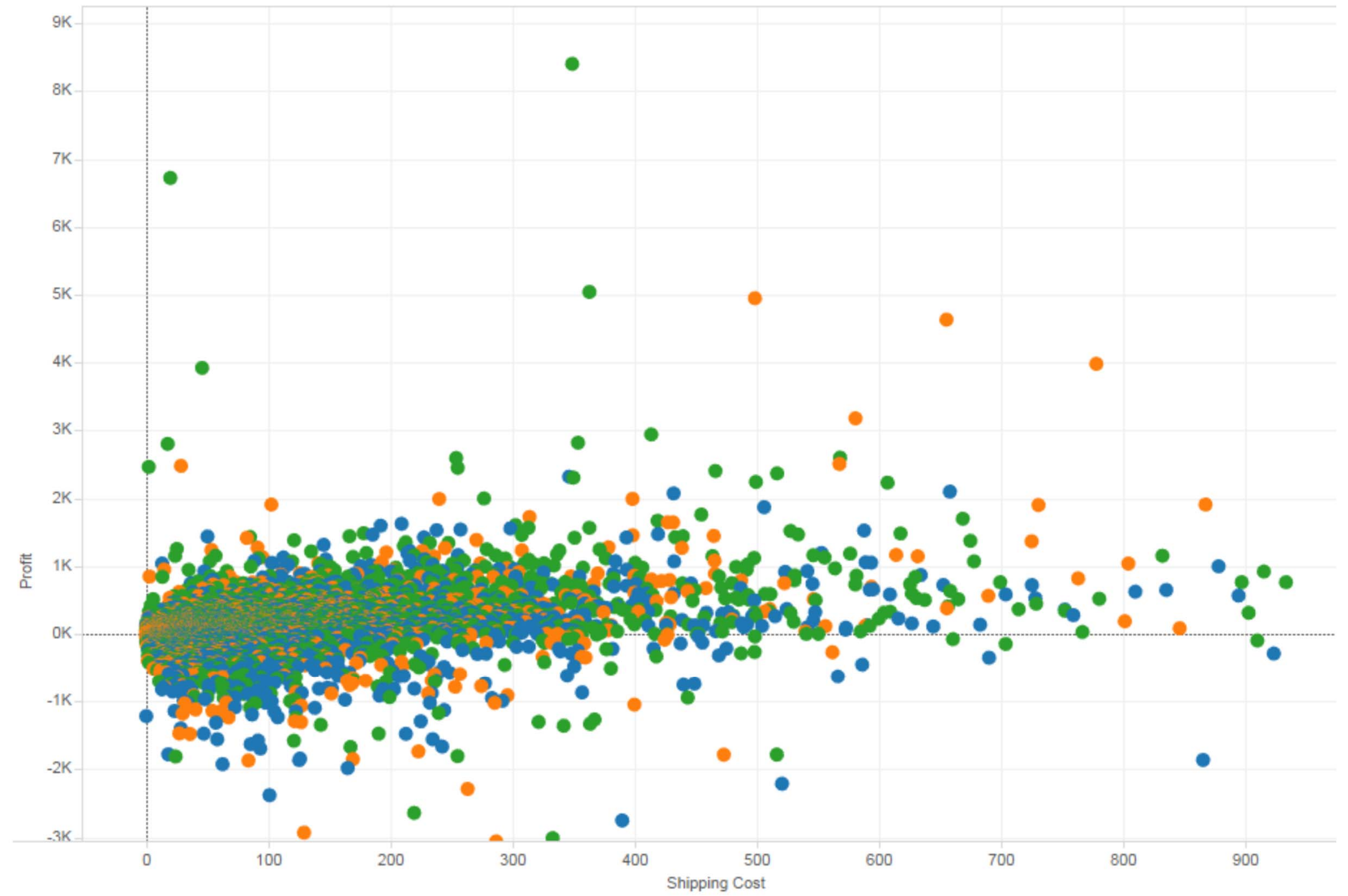
- Characterizing distributions
 - Outliers; extremes
- Finding correlations between two attributes

scale Items in the hundreds

→ Distinguishing data points from each other becomes difficult in larger data sets



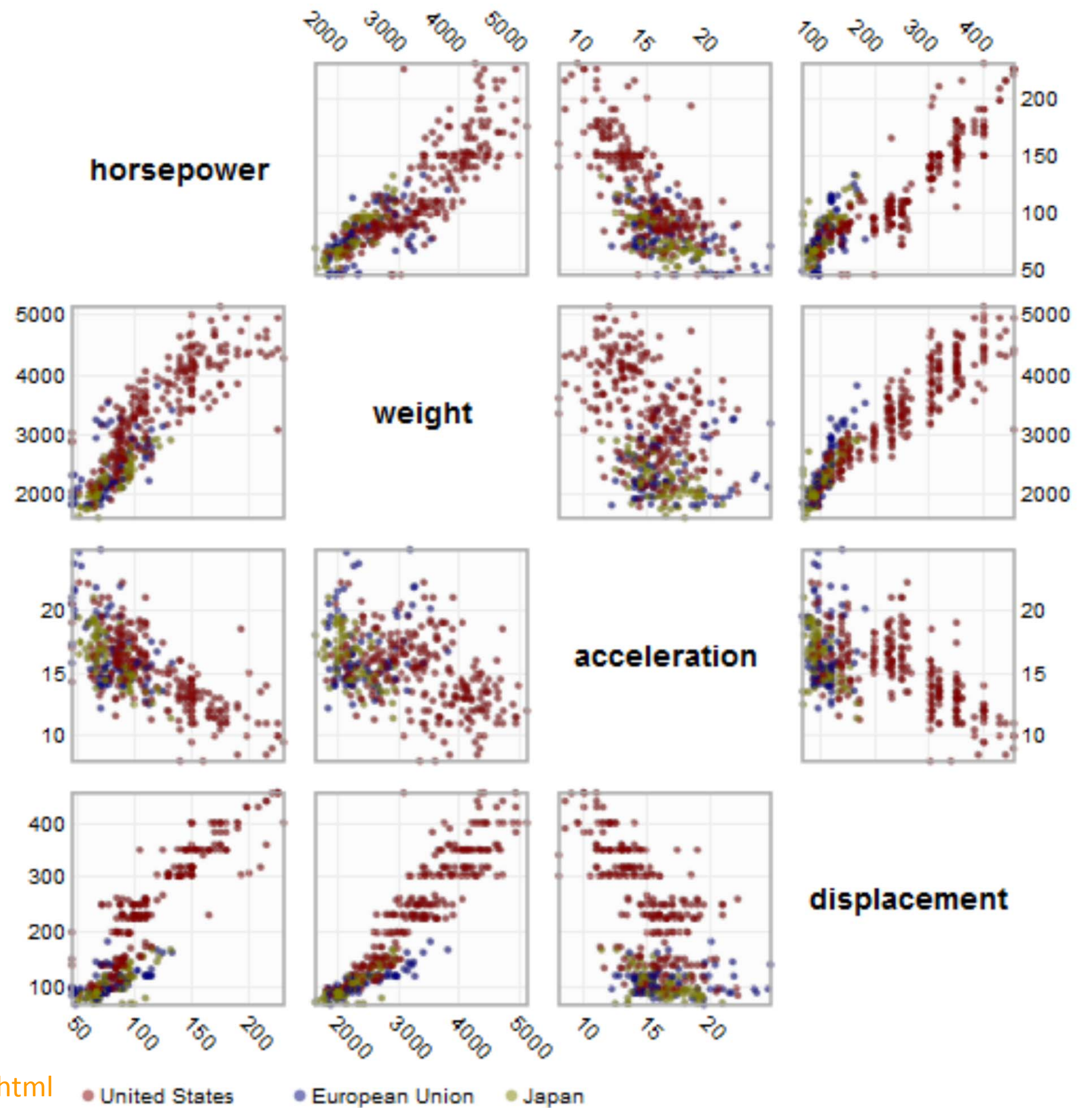
scatterplot



What if we want explore pairwise correlations across multiple quantitative attributes?

horsepower	weight	acceleration	displacement
50	3000	10	400
65	4000	25	300
250	3500	15	270
125	2100	17	350
200	2300	12	125

scatterplot matrix



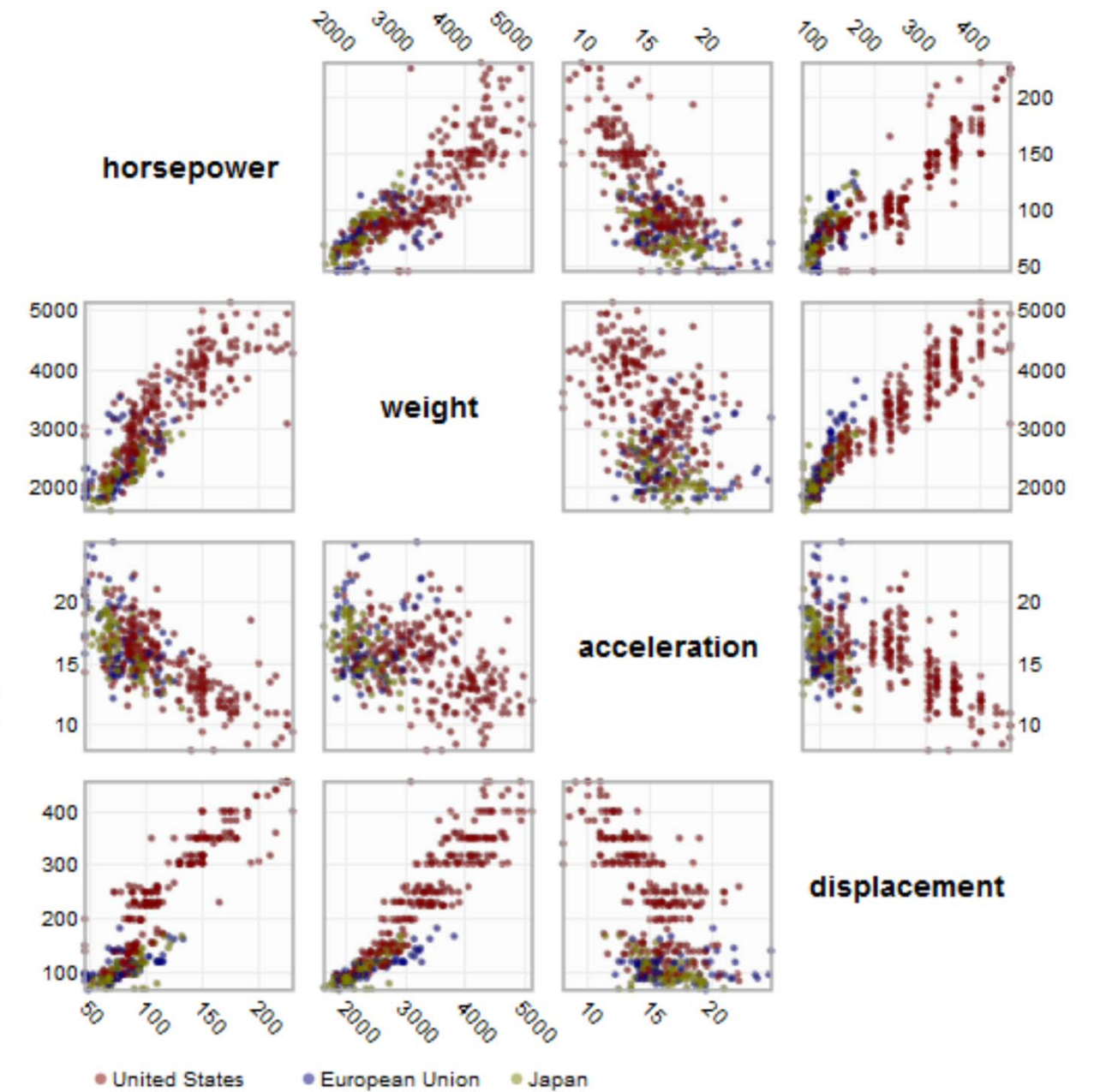
scatterplot matrix

why Find pairwise correlations

- Trends
- Outliers

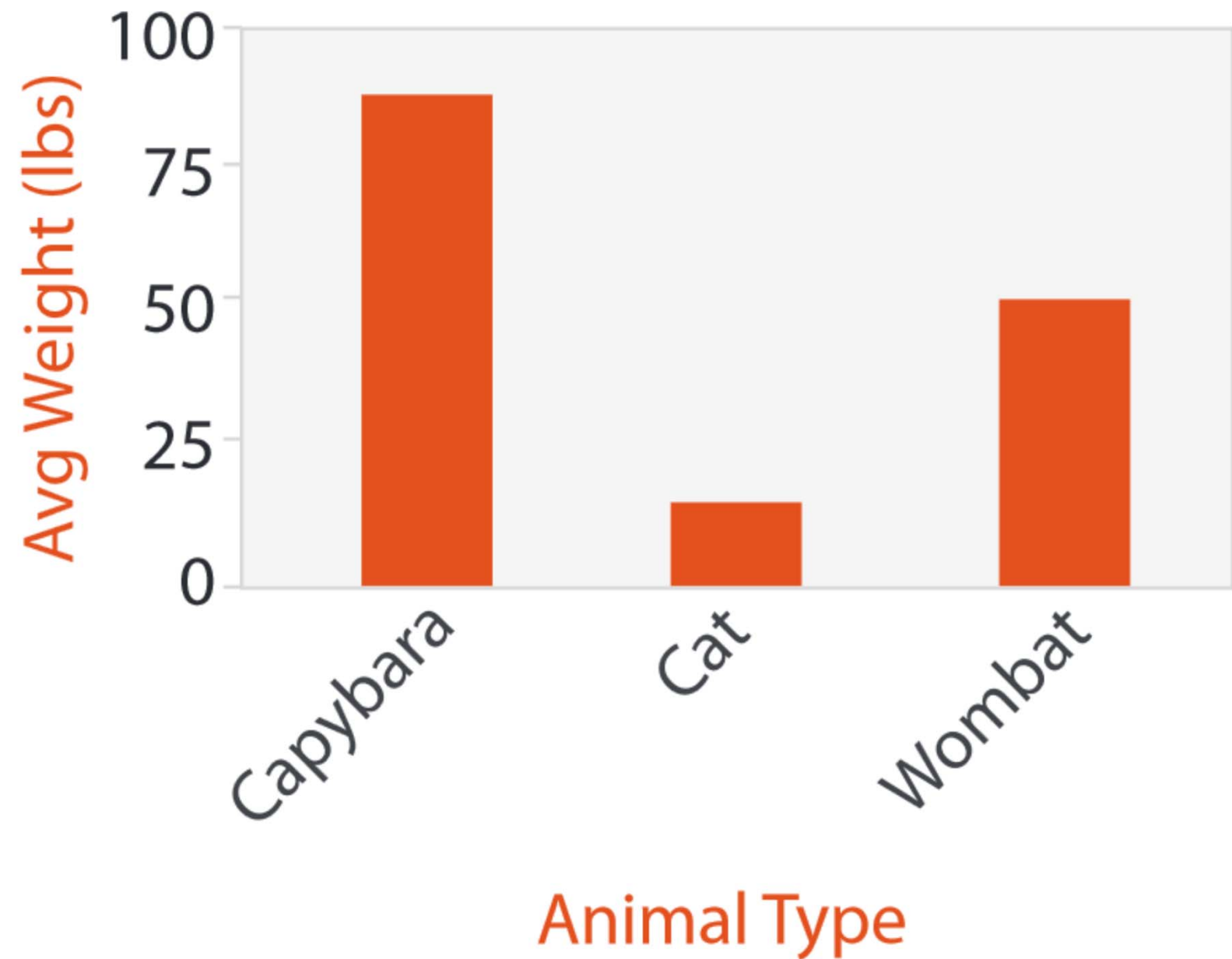
scale Attributes < 12

- Data points in the hundreds (as for scatterplots)



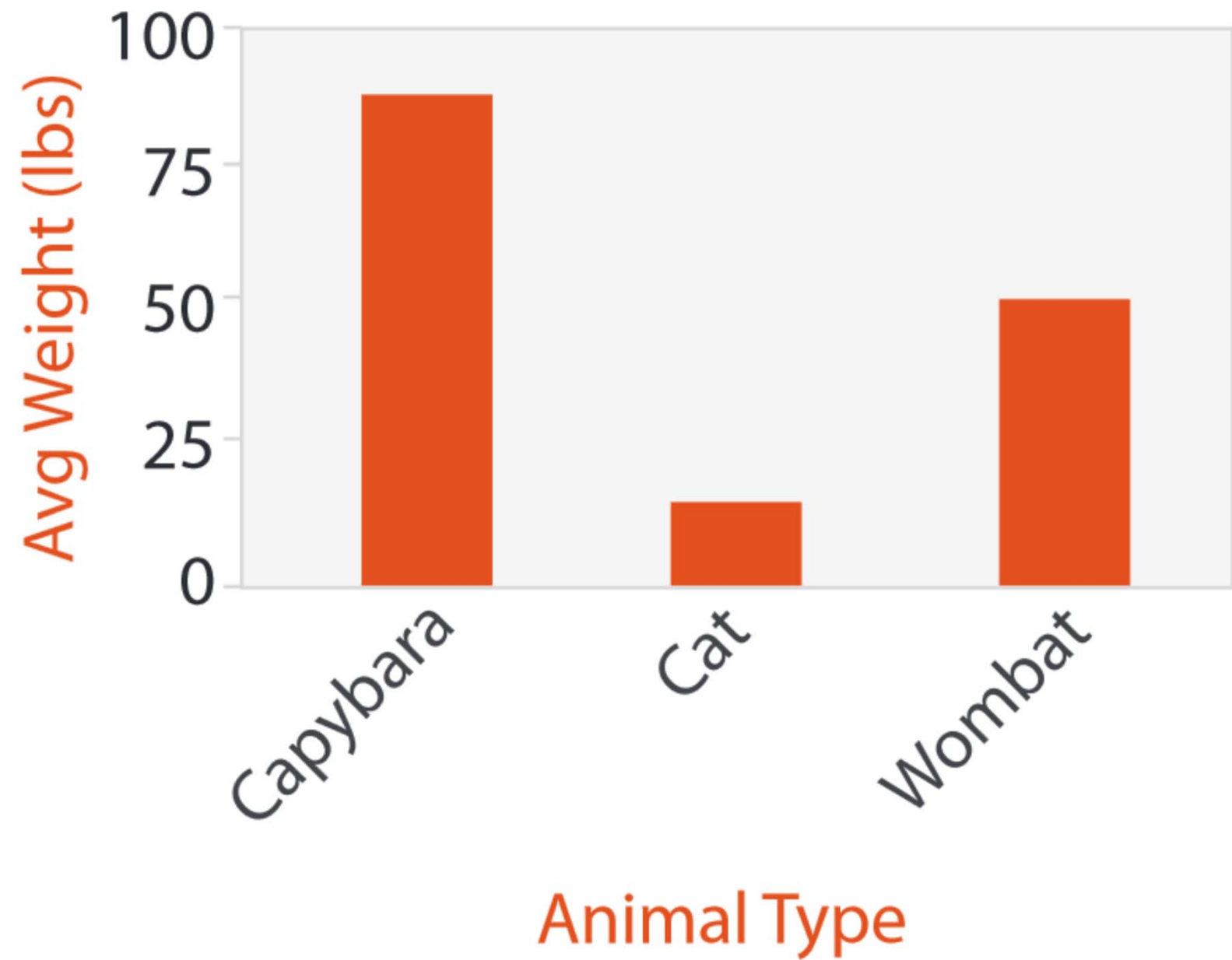
bar charts

ID	Animal type	Avg. weight
1	Capybara	78
2	Cat	13
3	Wombat	50



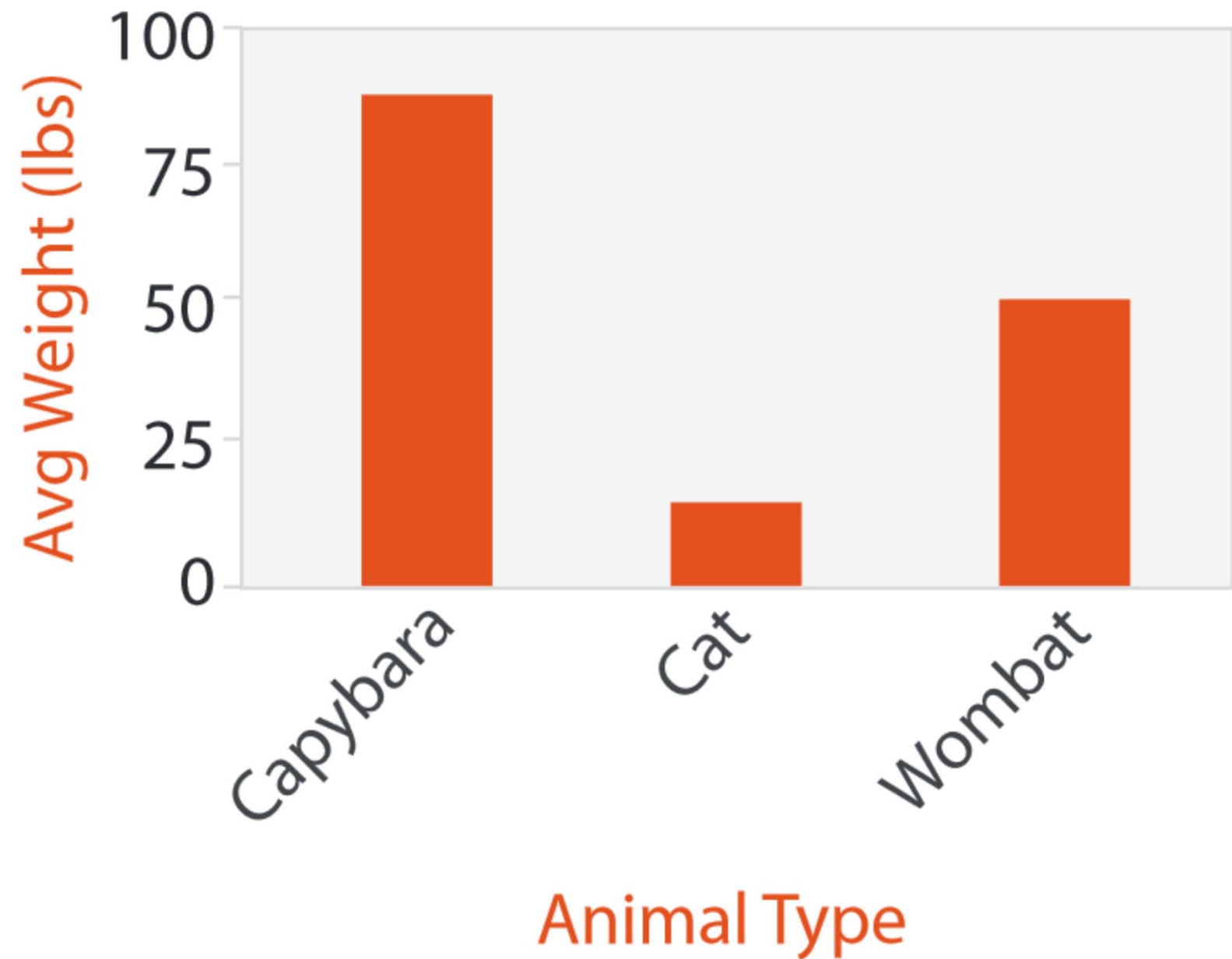
bar charts

ID	Animal type	Avg. weight
1	Capybara	78
2	Cat	13
3	Wombat	50



bar charts

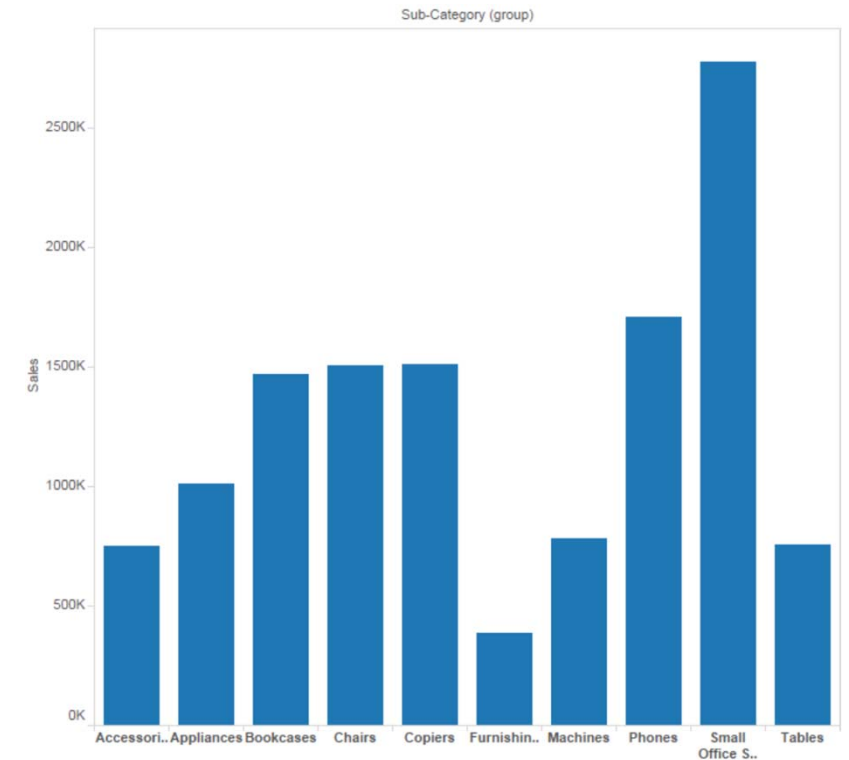
ID	Animal type	Avg. weight
1	Capybara	78
2	Cat	13
3	Wombat	50



bar charts

what → One categorical attribute and one quantitative attribute

- Line marks



how → Value attribute expressed through **length** and **aligned in vertical position**

- Levels of nominal dimension separated through **horizontal position**

why → Lookup

- Compare values
- Presenting trends

bar charts



<http://www.oecdbetterlifeindex.org/#/11111111111>

bar charts

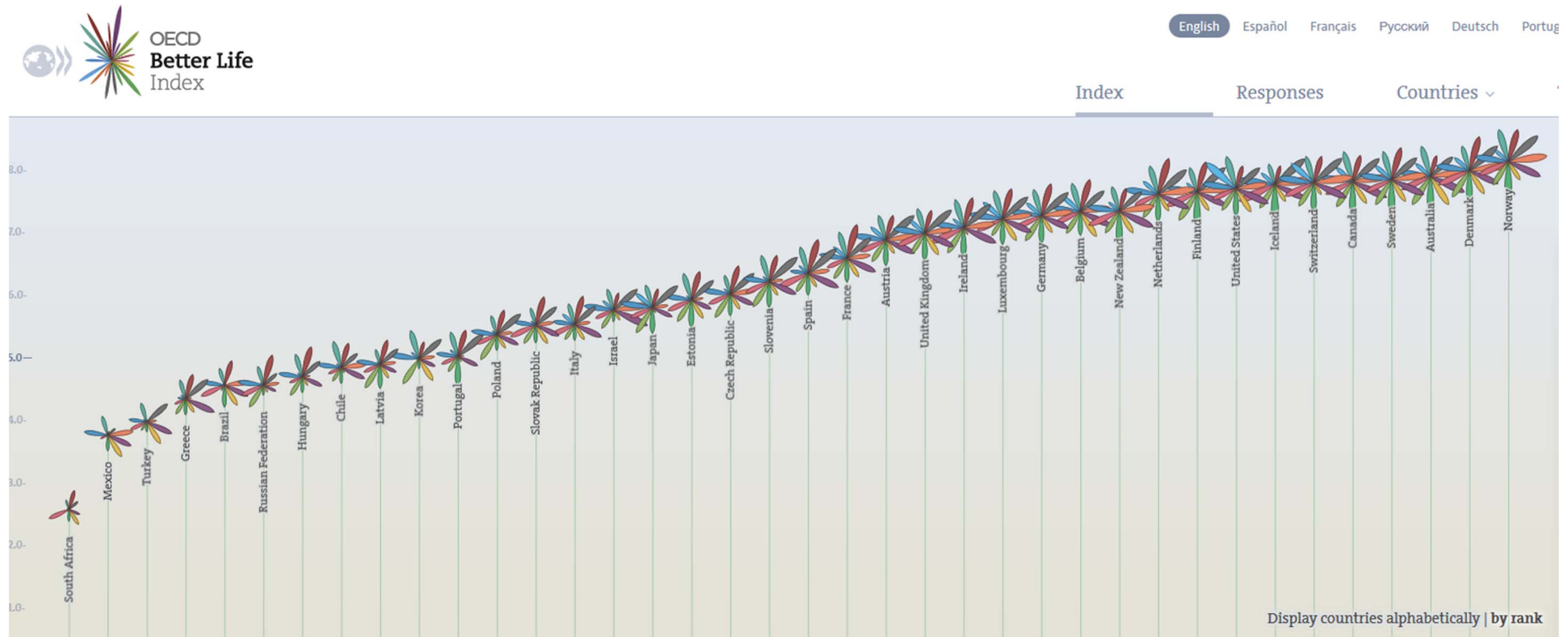
- Alphabetical ordering
 - Good for lookup



<http://www.oecdbetterlifeindex.org/#/11111111111>

bar charts

- Quantity-driven ordering
 - Good for seeing trends

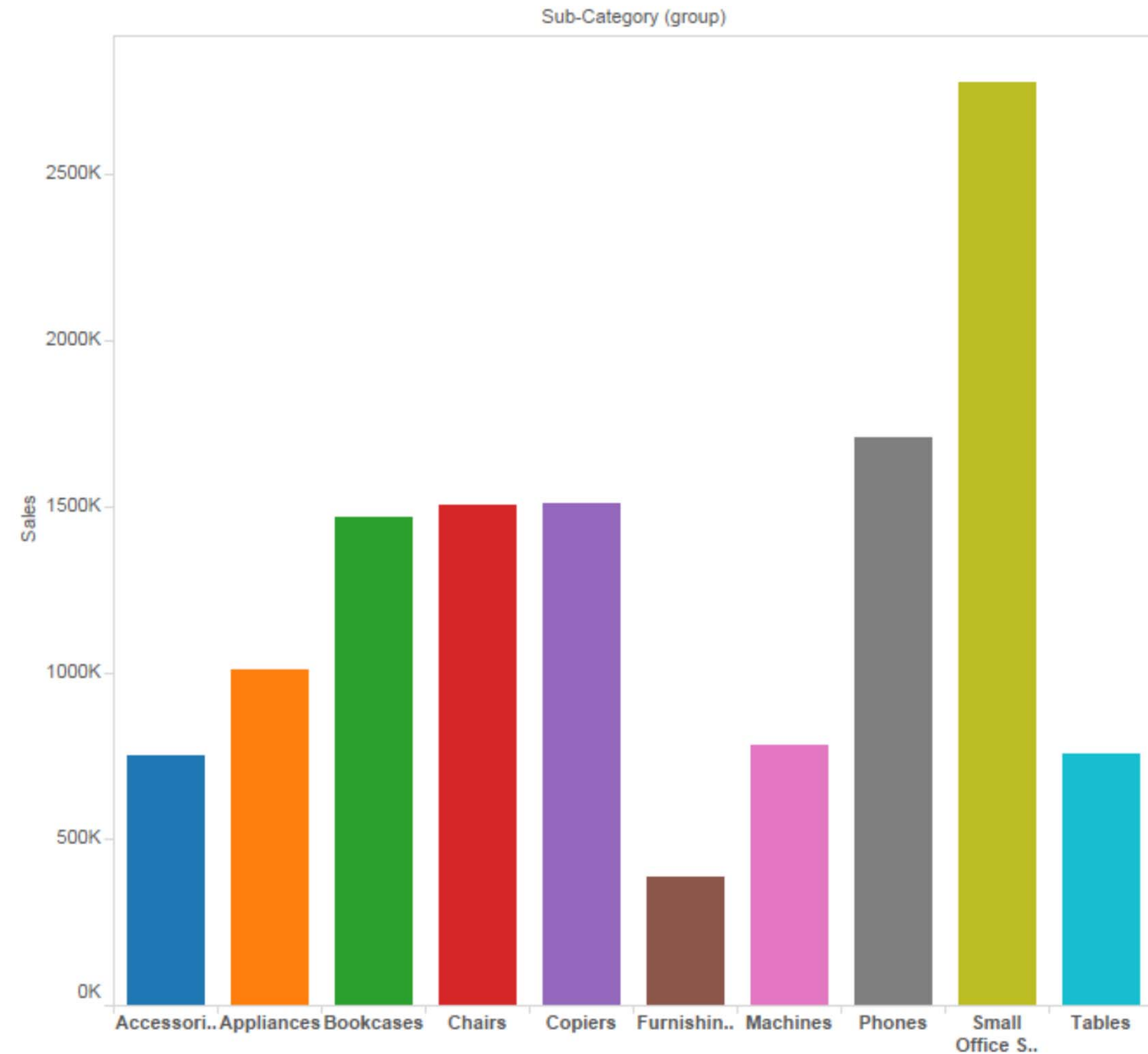


<http://www.oecdbetterlifeindex.org/#/1111111111>

bar charts

double-encoding

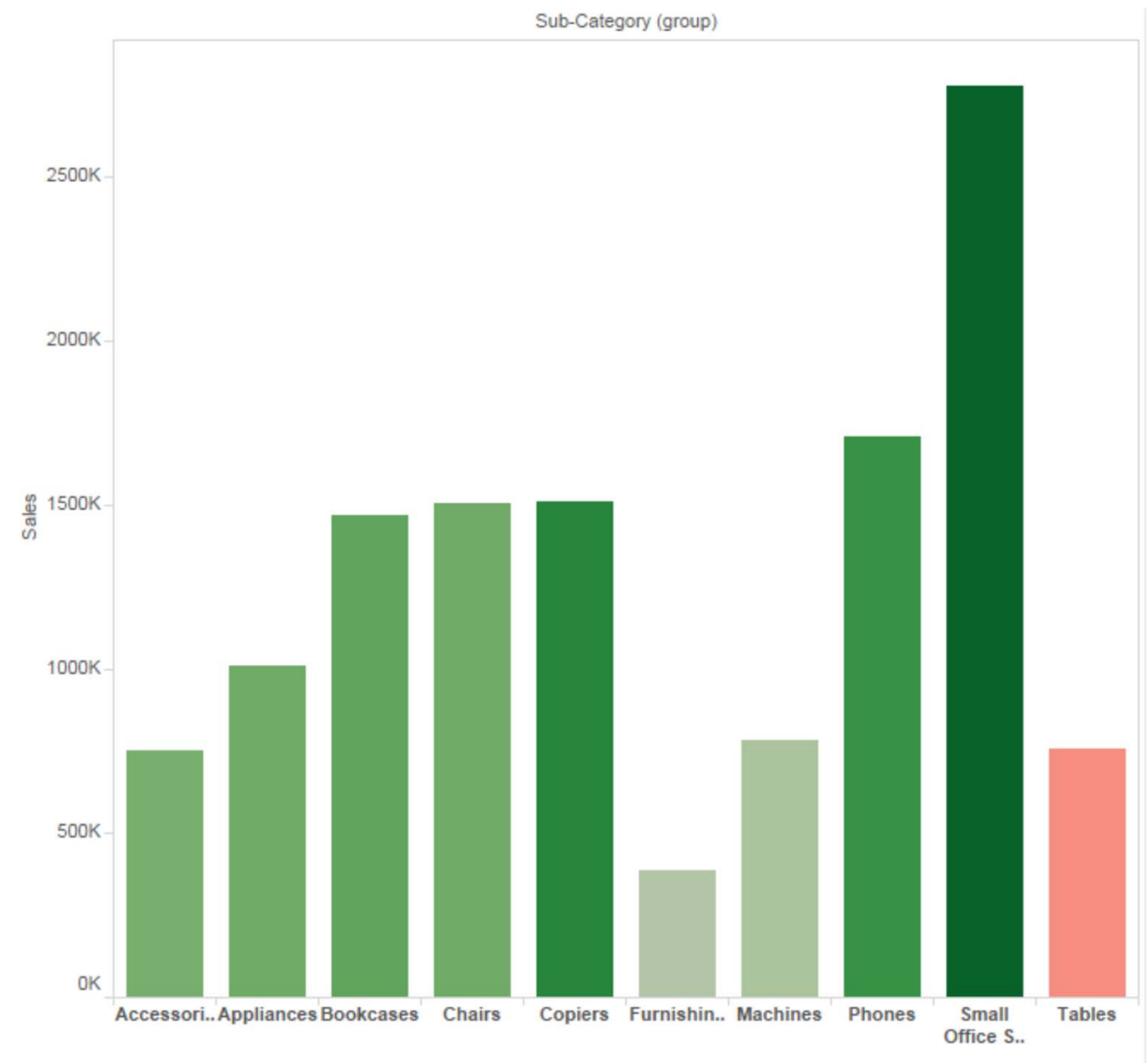
- Product category
 - x-position
 - Colour (hue)
- Sales
 - Length



bar charts

one categorical + two quantitative attributes

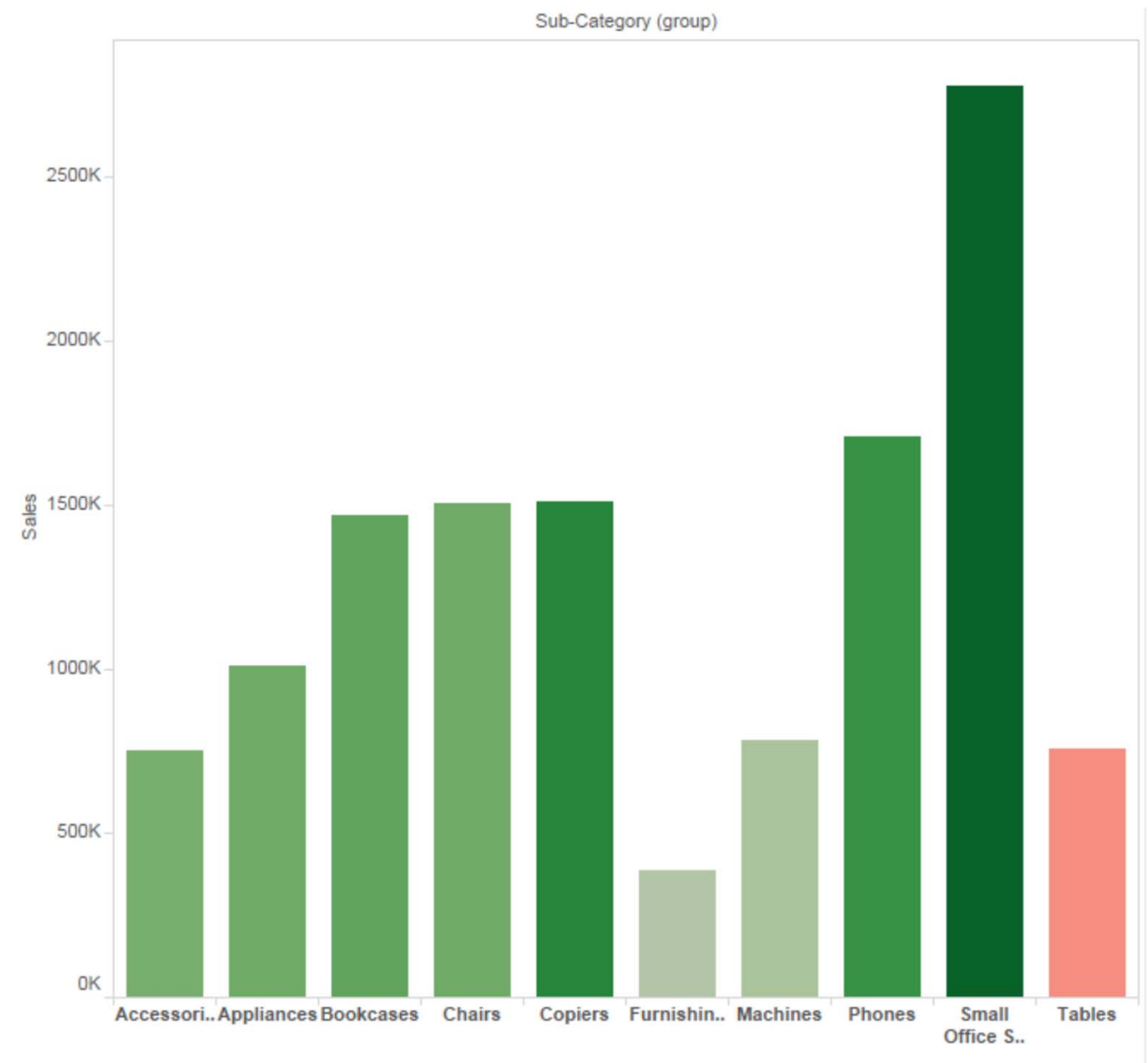
- Product category
 - x-position
 - Sales
 - Length
 - Profit amount
 - Colour (saturation)
 - Profit category
 - Colour (hue)
 - Red: minus
 - Green: plus
- Diverging colour scale



bar charts

one categorical + two quantitative attributes

- Product category
 - x-position
 - Sales
 - Length
 - Profit amount
 - Colour (saturation)
 - Profit category
 - Colour (hue)
 - Red: minus
 - Green: plus
- Diverging colour scale



bar charts

what → One categorical and one quantitative attribute

- Line marks

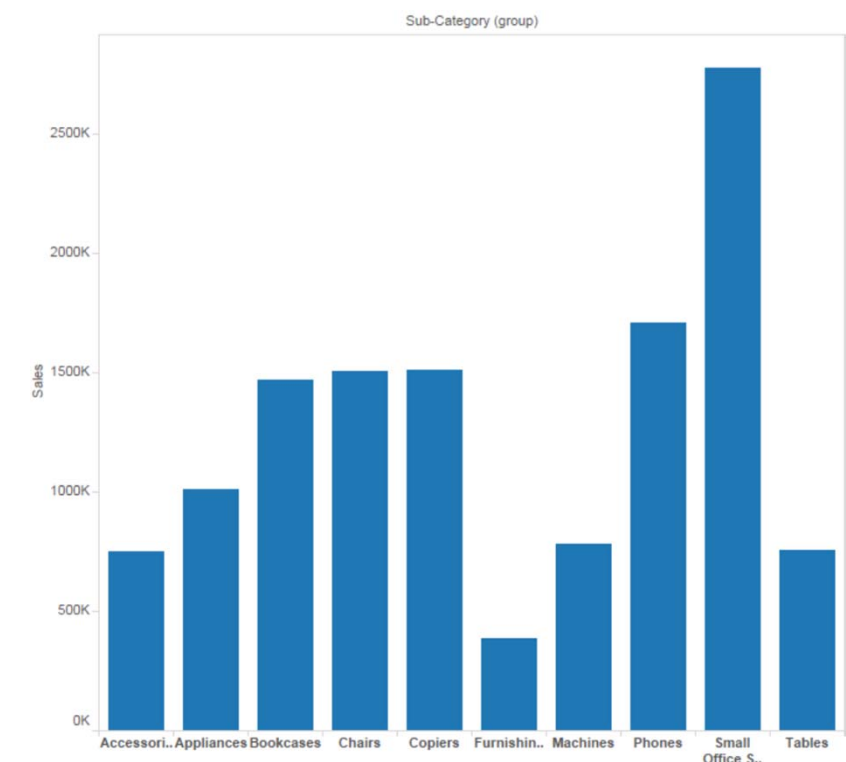
how → Value attribute expressed through length and aligned in vertical position

- Levels of nominal dimension separated through horizontal position

why → Lookup

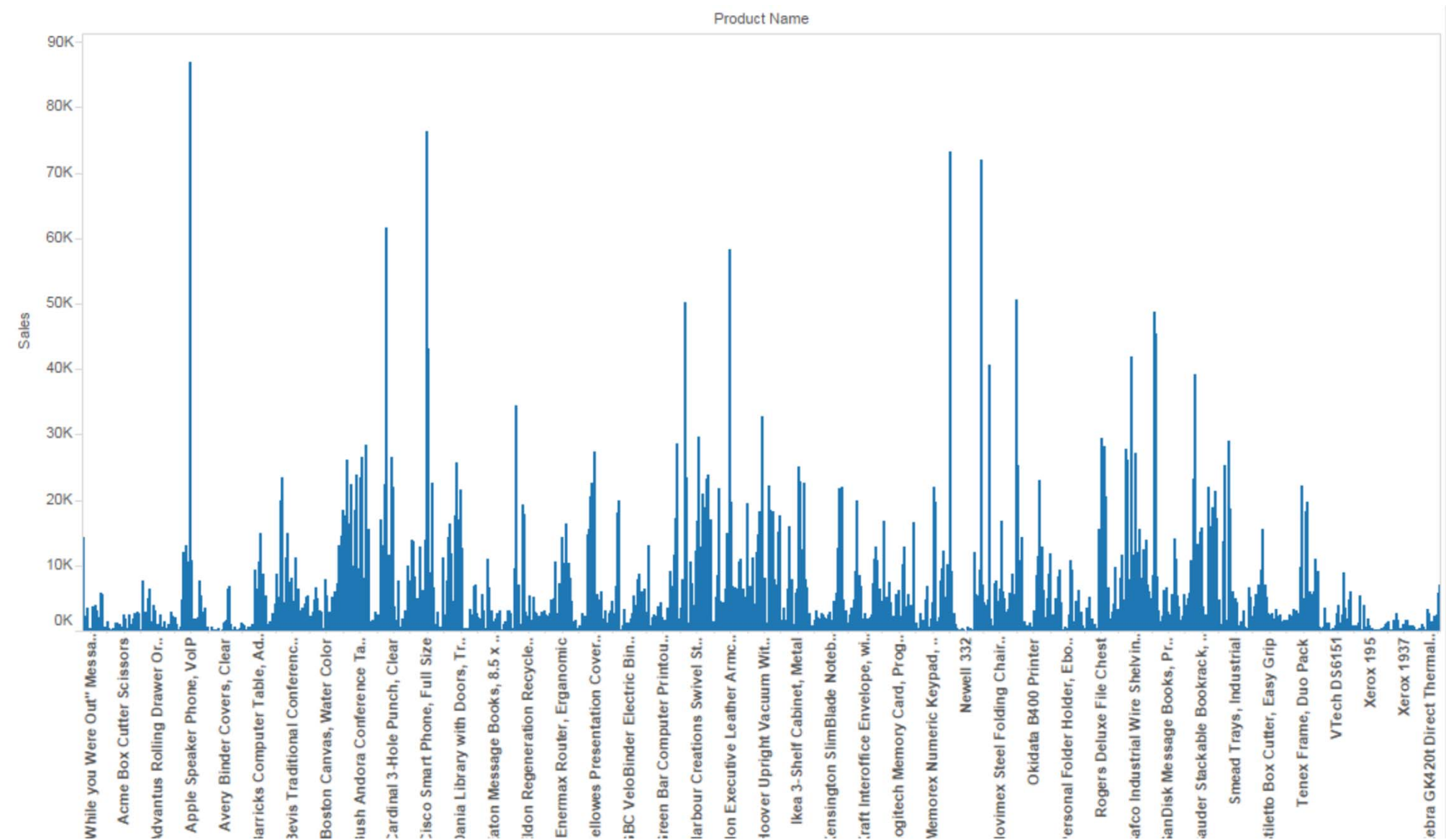
- Compare values
- Presenting trends

scale → Limited to hundreds of key levels



bar charts

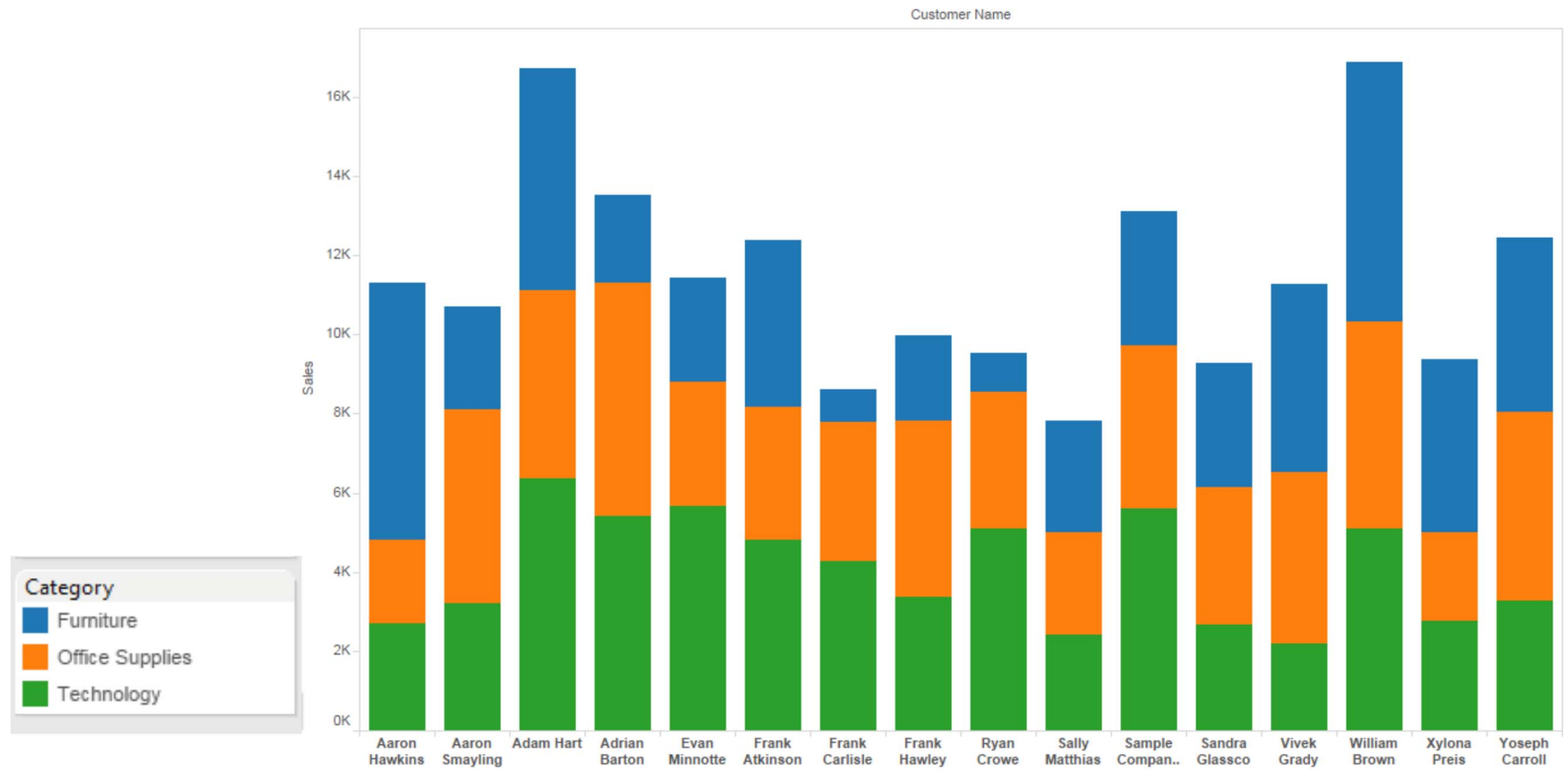
- Over 3000 levels (unique categories)



bar charts & multidimensional data

- Multidimensional table
 - One categorical attribute can be divided into several sub-categories
- Example:
 - Data on different individual customers buys different types of goods:
 - Furniture
 - Office supplies
 - Technology
 - We want to show the total of sales per customer and category of goods

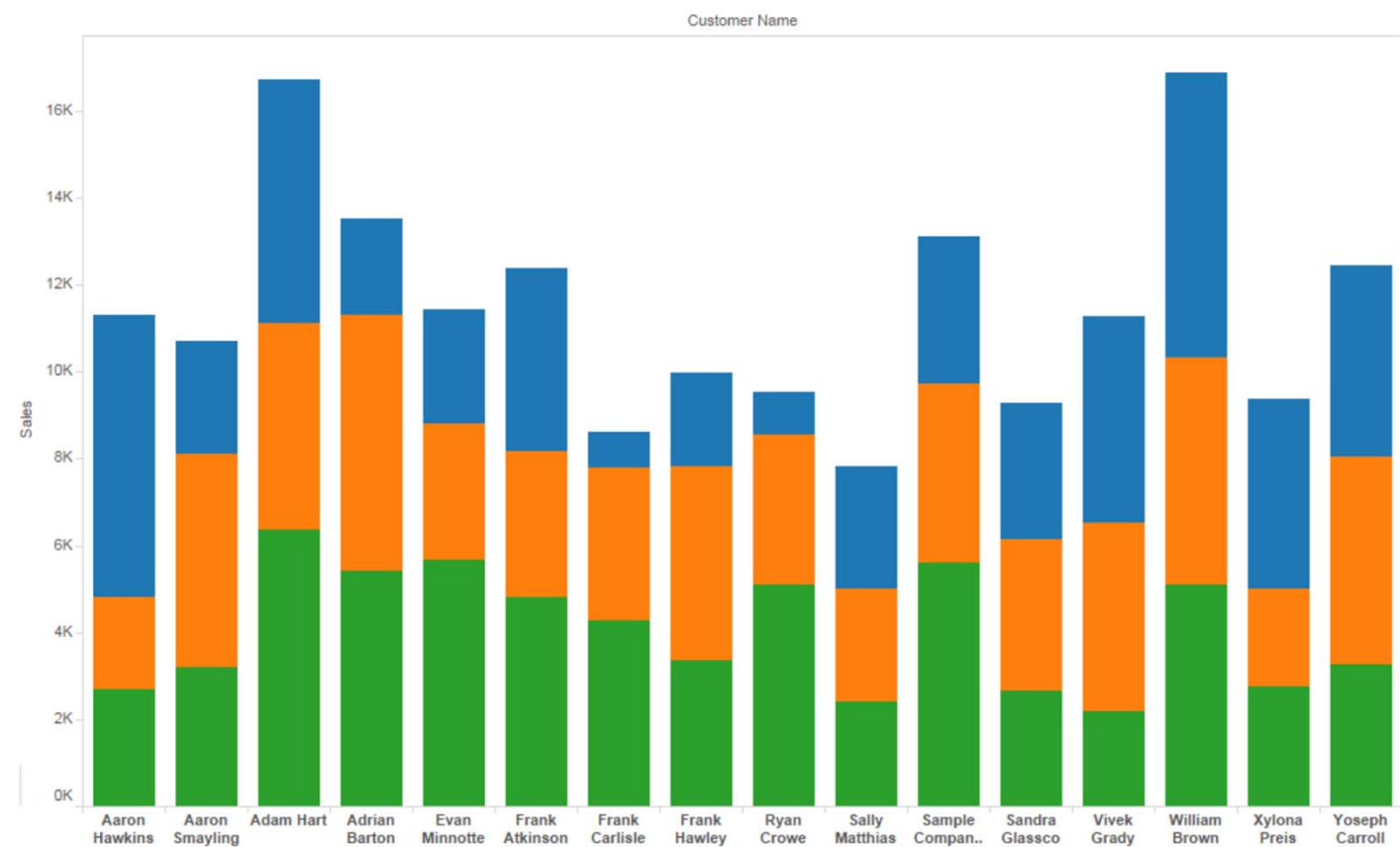
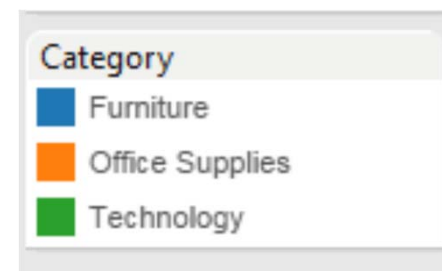
stacked bar charts



stacked bar charts

what 2 categorical attributes; 1 quantitative attribute

- Customer (primary category)
- Product category (secondary category)
- Sale (quantitative attribute)



stacked bar charts

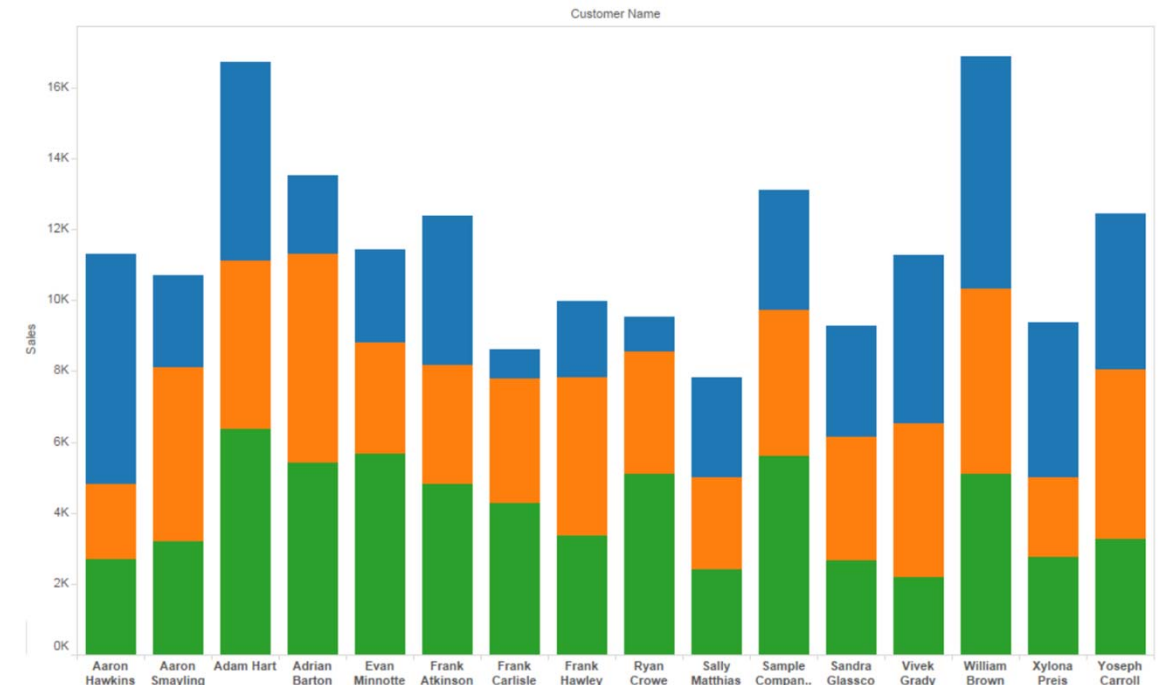
what 2 categorical attributes; 1 quantitative attribute

- Customer (primary category)
- Product category (secondary category)
- Sale (quantitative attribute)

how Bar glyphs where **lengths** represents the value of each category of the secondary category

→ Full bar height shows sum of all subcategories

- Primary category levels separated through **horizontal position**
- Secondary category levels separated through **colour (hue)**



stacked bar charts

why

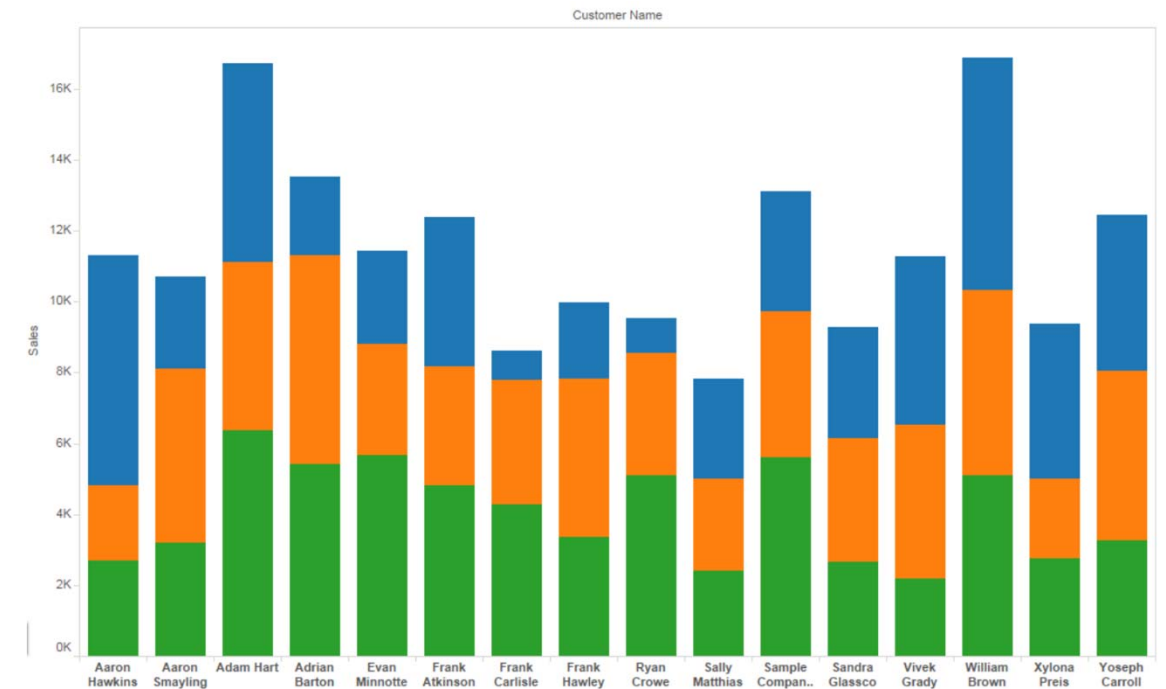
Part-of-a-whole relationship using absolute values

- Lookup values
- Find trends
- Limitations
 - Height of full bars and lowest sub-bars easy to compare
 - Comparison of stacked sub-bars more difficult: no shared baseline
 - Stacking order is important; interactivity can help

scale

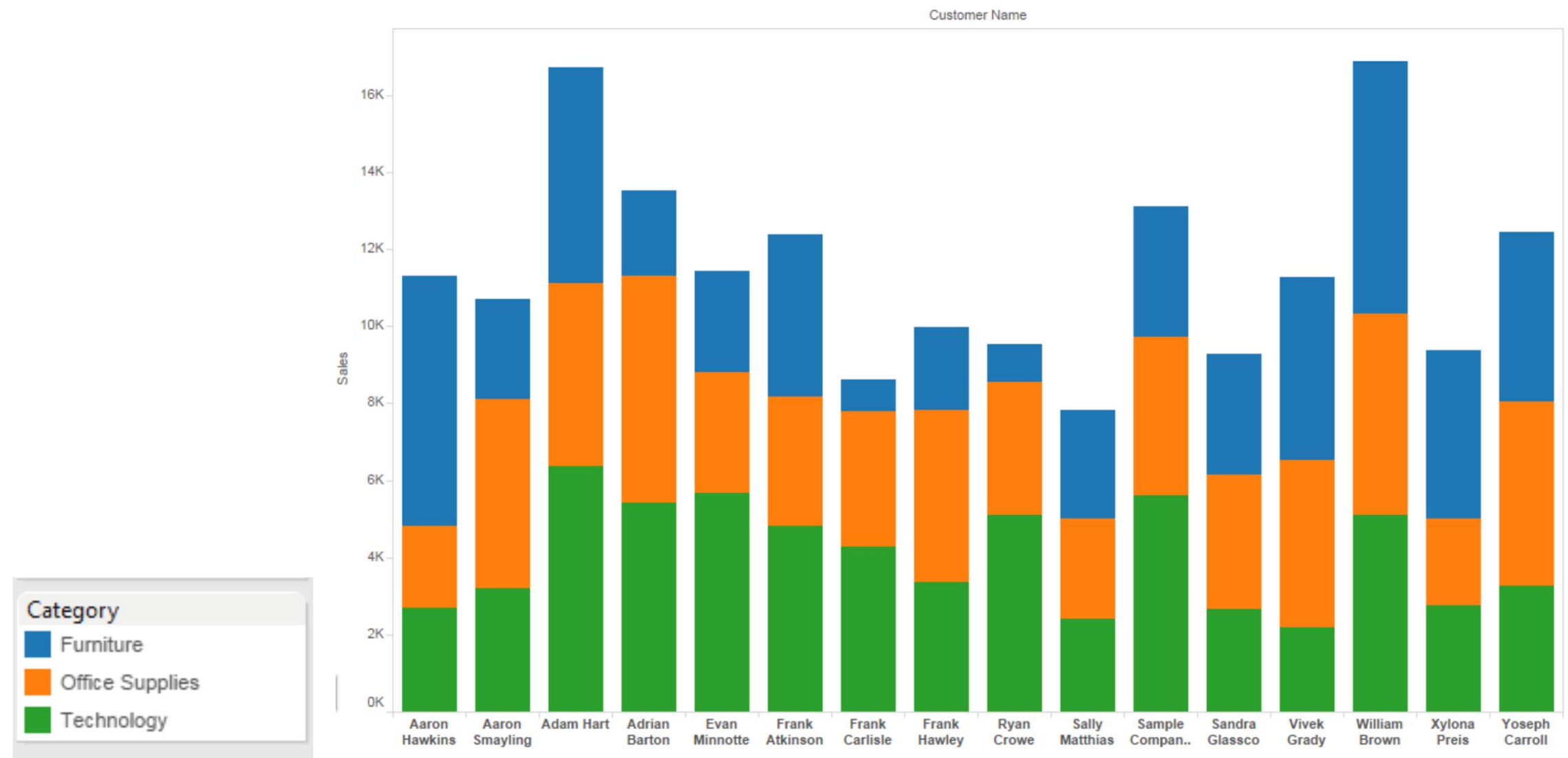
Scalability:

- Limited to hundreds of primary category levels
- Limited to several to a dozen secondary category levels



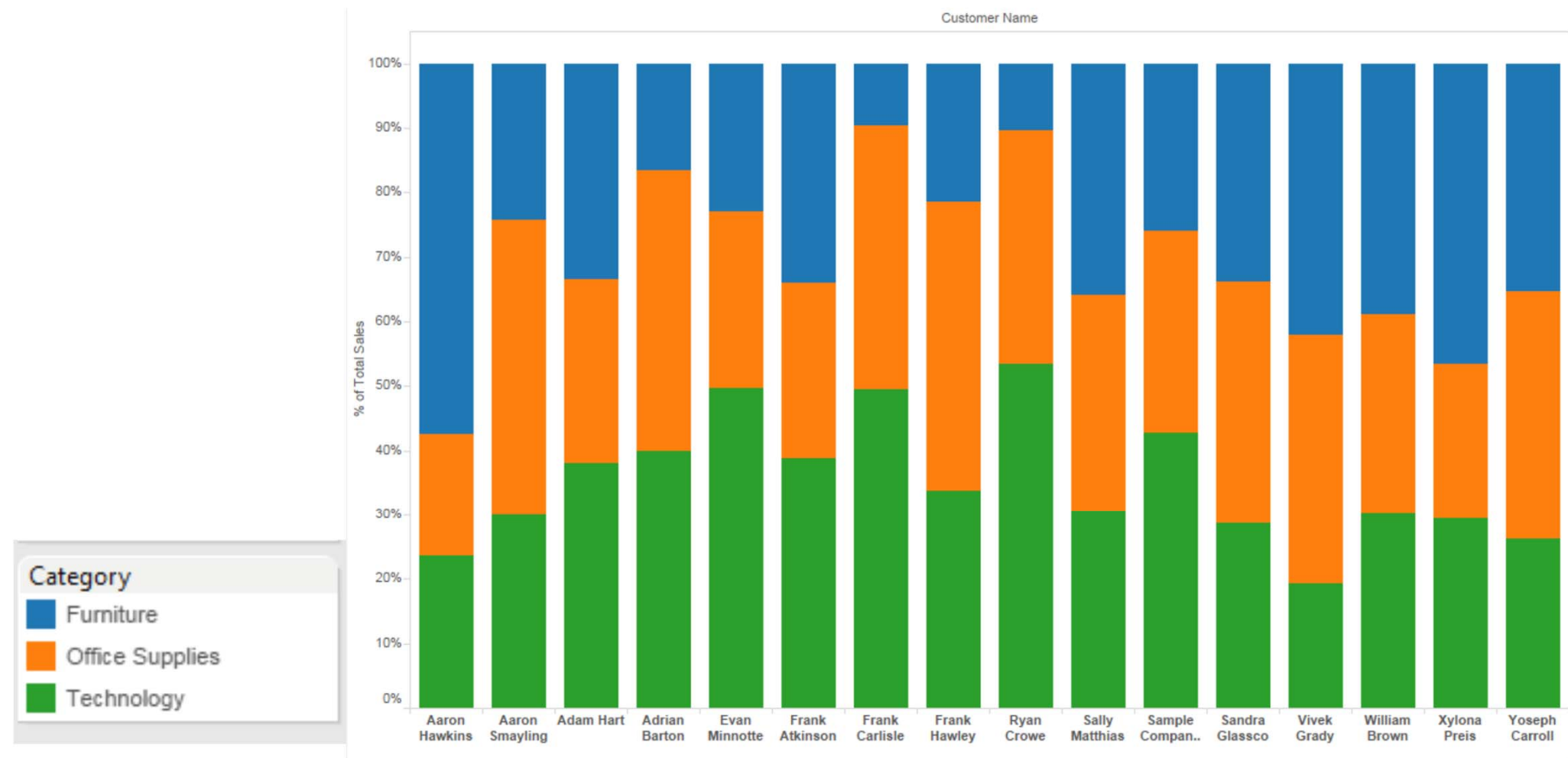
stacked bar graphs

- Typical: absolute values



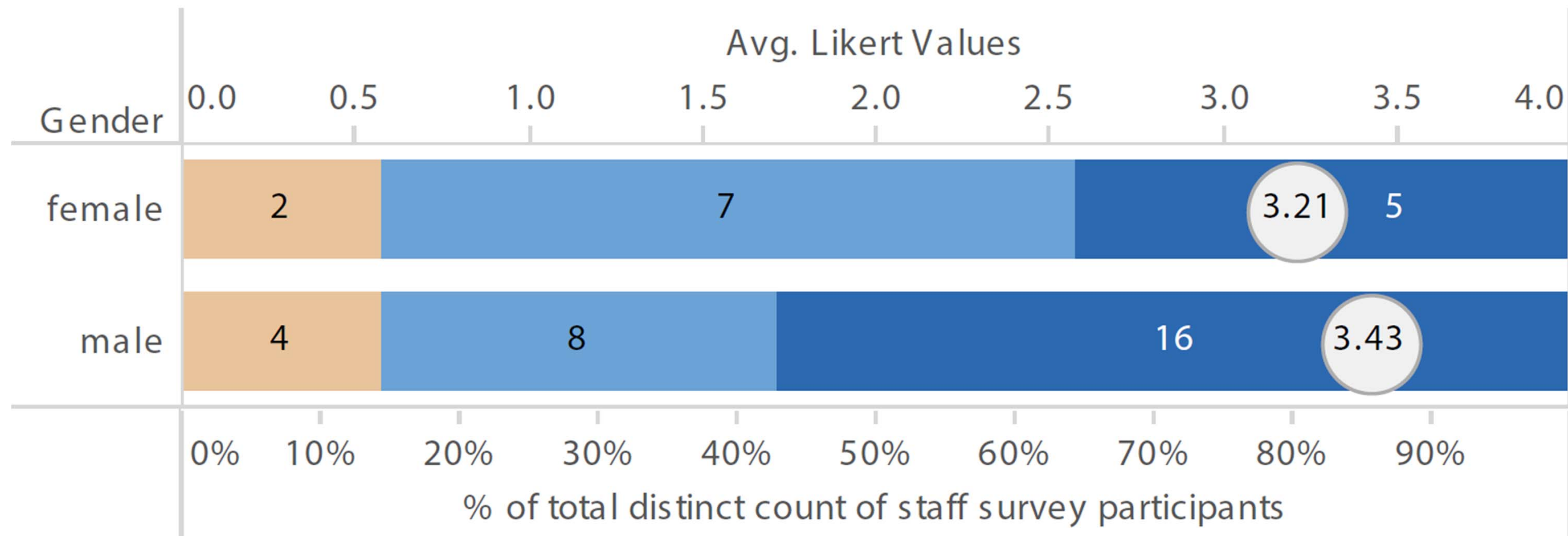
normalized stacked bar charts

- Showing relative proportions
- Part-of-whole relationships; comparison of distribution across key attribute levels

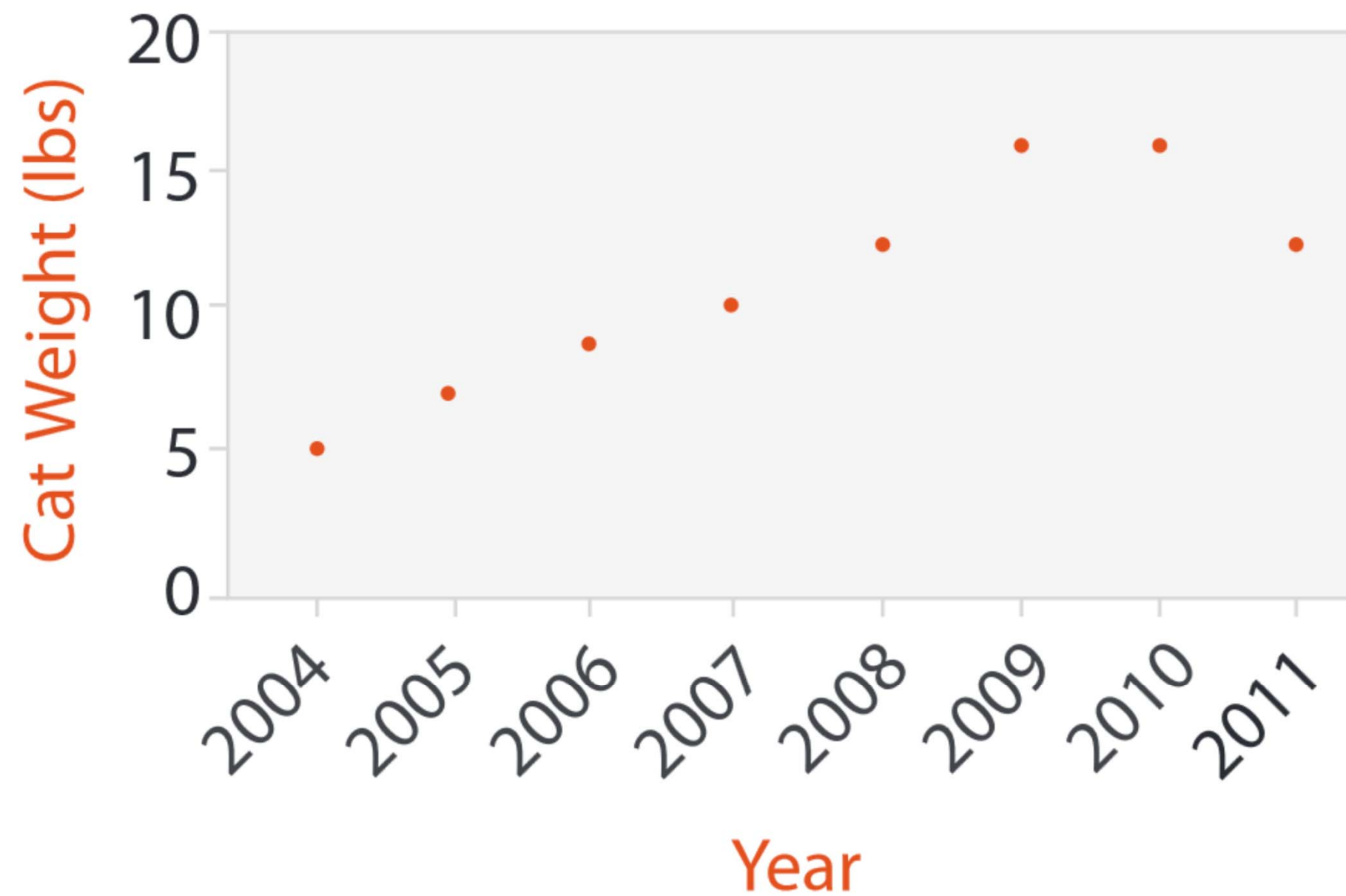


normalized stacked bar charts

- Showing relative proportions
- Part-of-whole relationships; comparison across category attribute levels



dot chart / dot plot



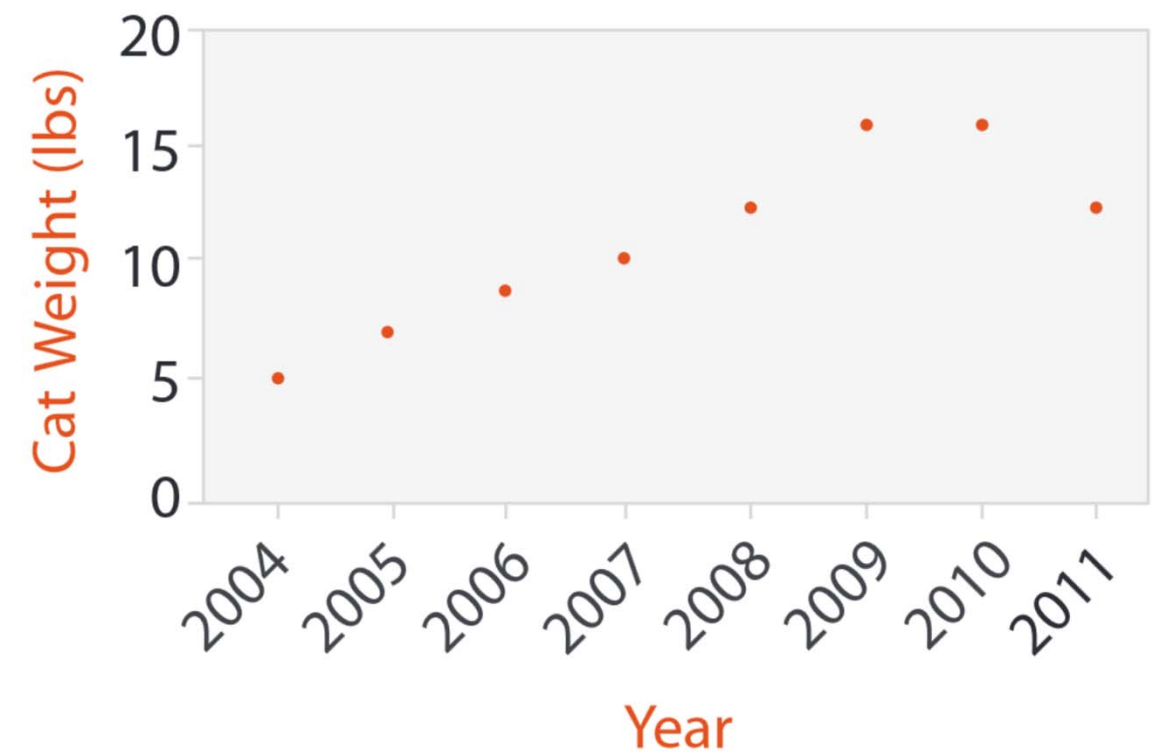
dot chart / dot plot

what Difference to scatter plots

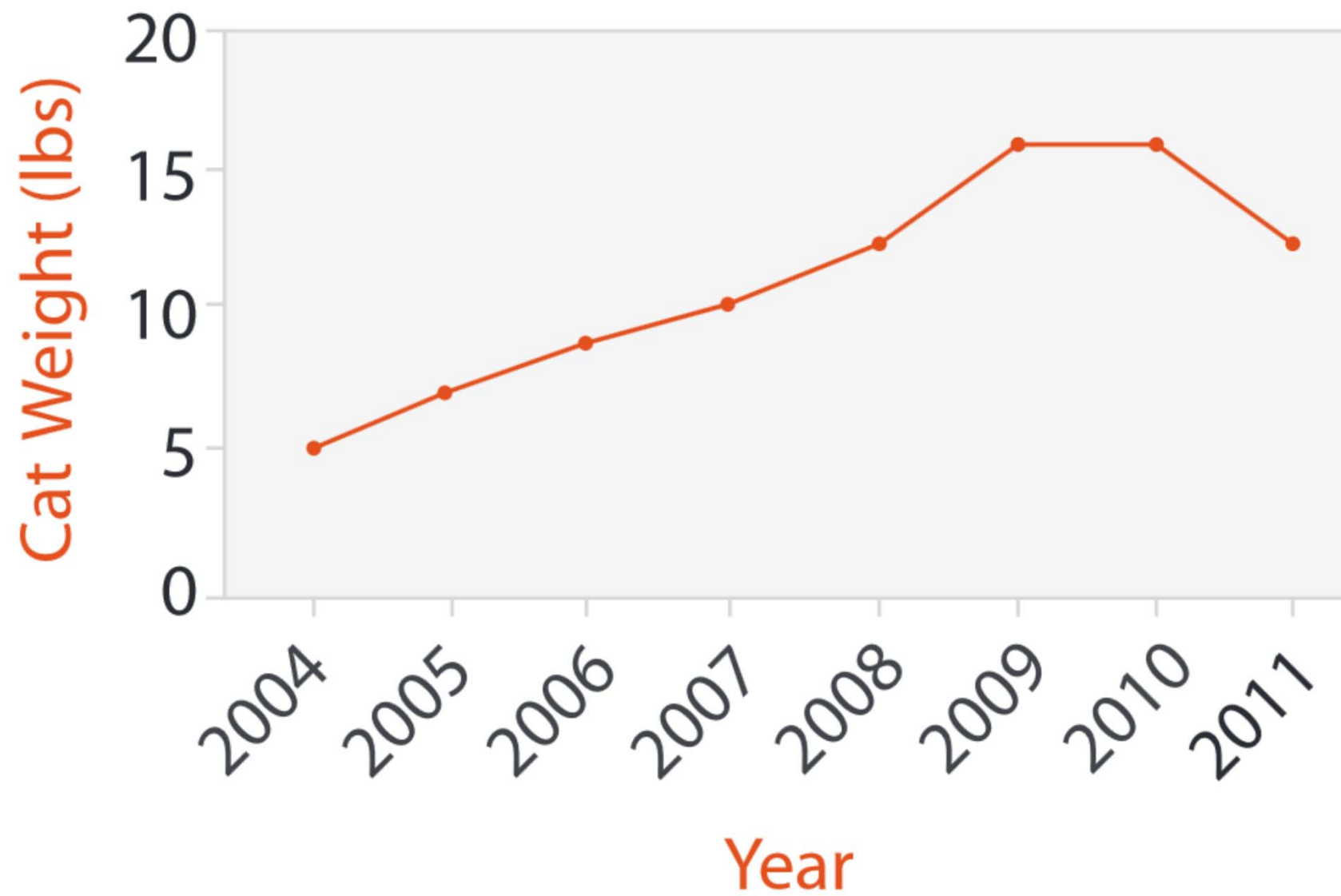
- One **quantitative** and one **ordinal attribute** instead of two quantitative attributes

how Difference to bar charts

- Point mark instead of line mark
- Ordinal instead of categorical
- Dots can imply connection and/or order



line chart / line graph

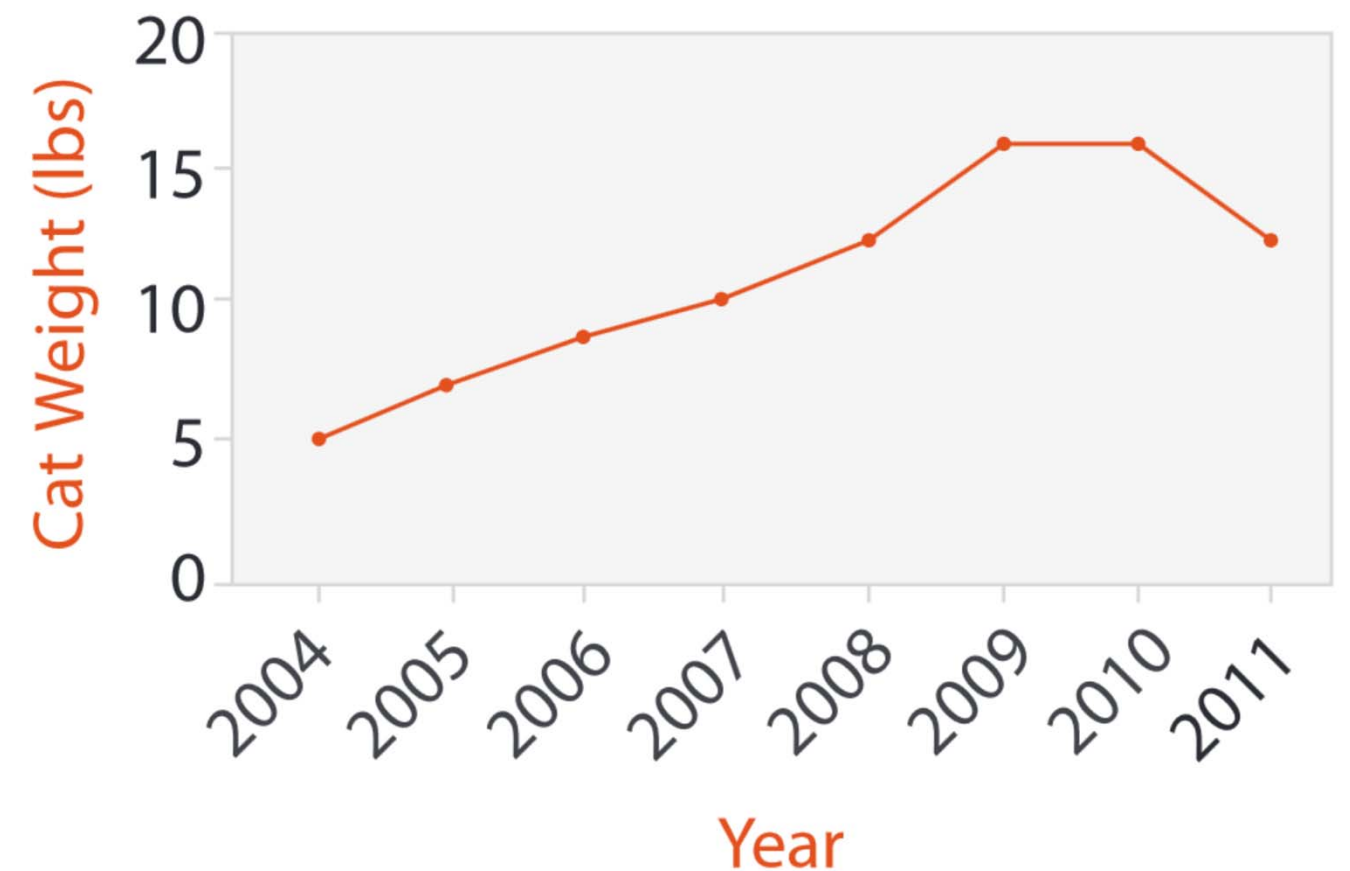


line chart / line graph

- Follow directly from dot charts with a line connecting the different data points

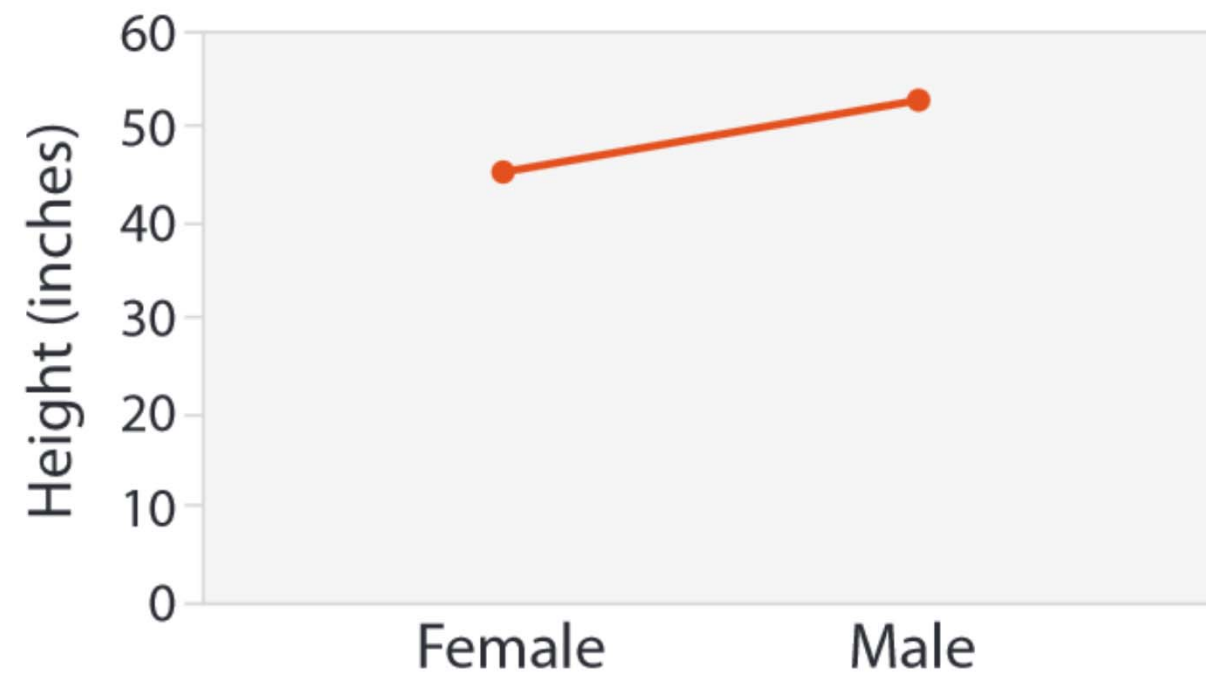
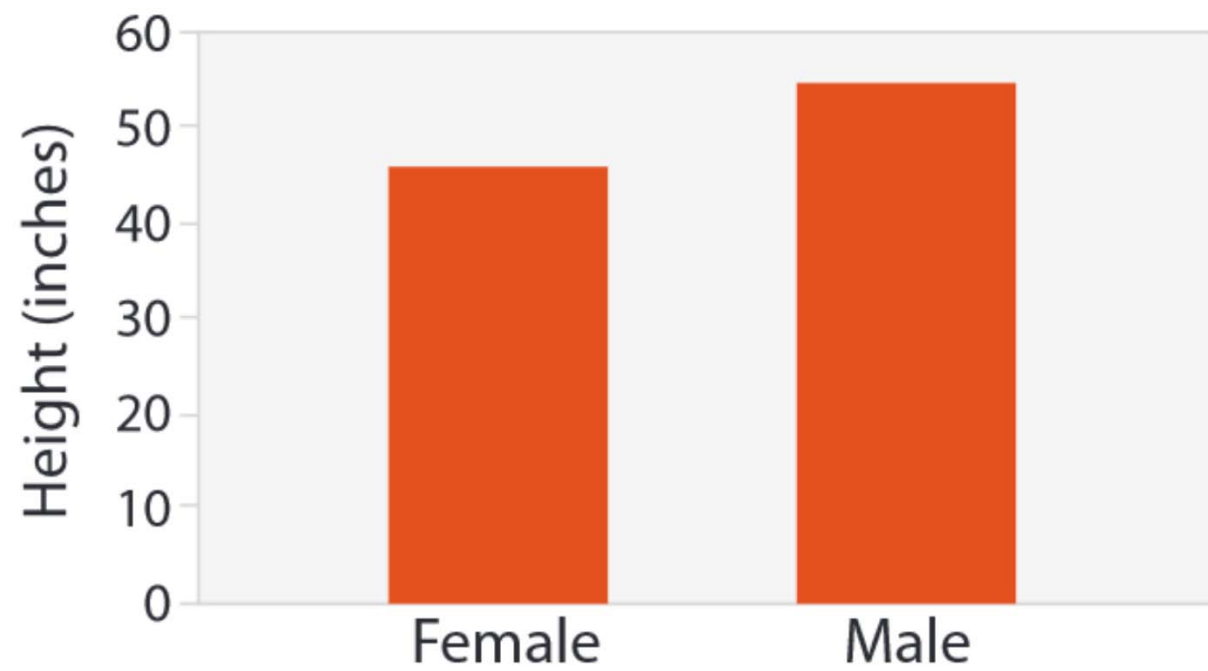
why

Good for showing trends



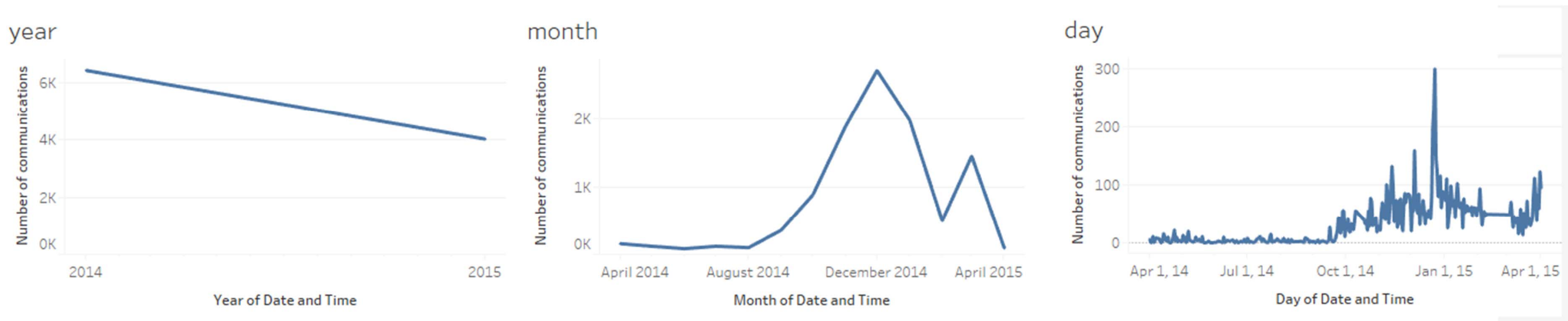
line chart / line graph

- Line emphasizes ordering & continuity
 - Do not use for categorical data!!



line chart / line graph

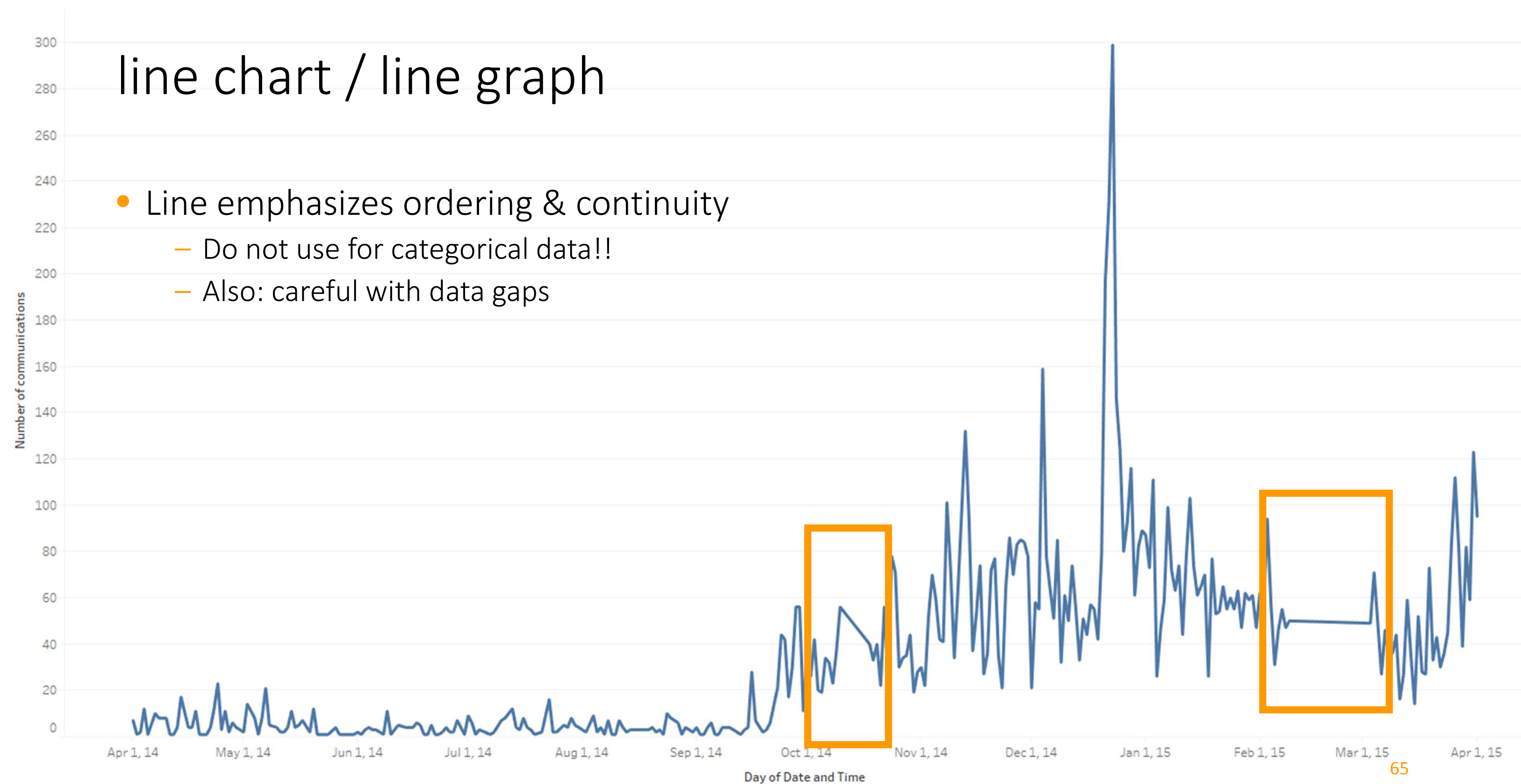
- Line emphasizes ordering & continuity
 - Do not use for categorical data!!
 - Also: careful with data gaps



day

line chart / line graph

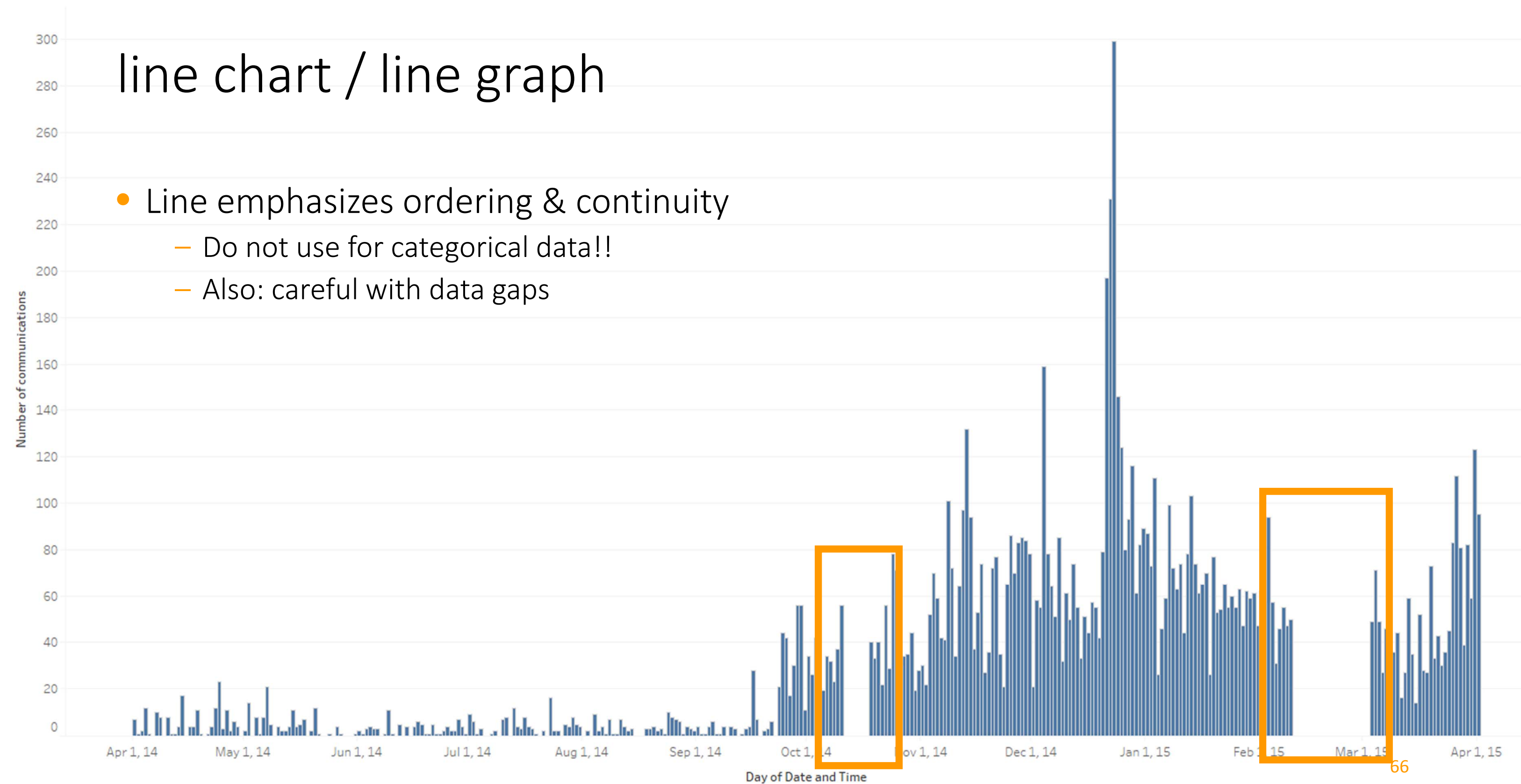
- Line emphasizes ordering & continuity
 - Do not use for categorical data!!
 - Also: careful with data gaps



day

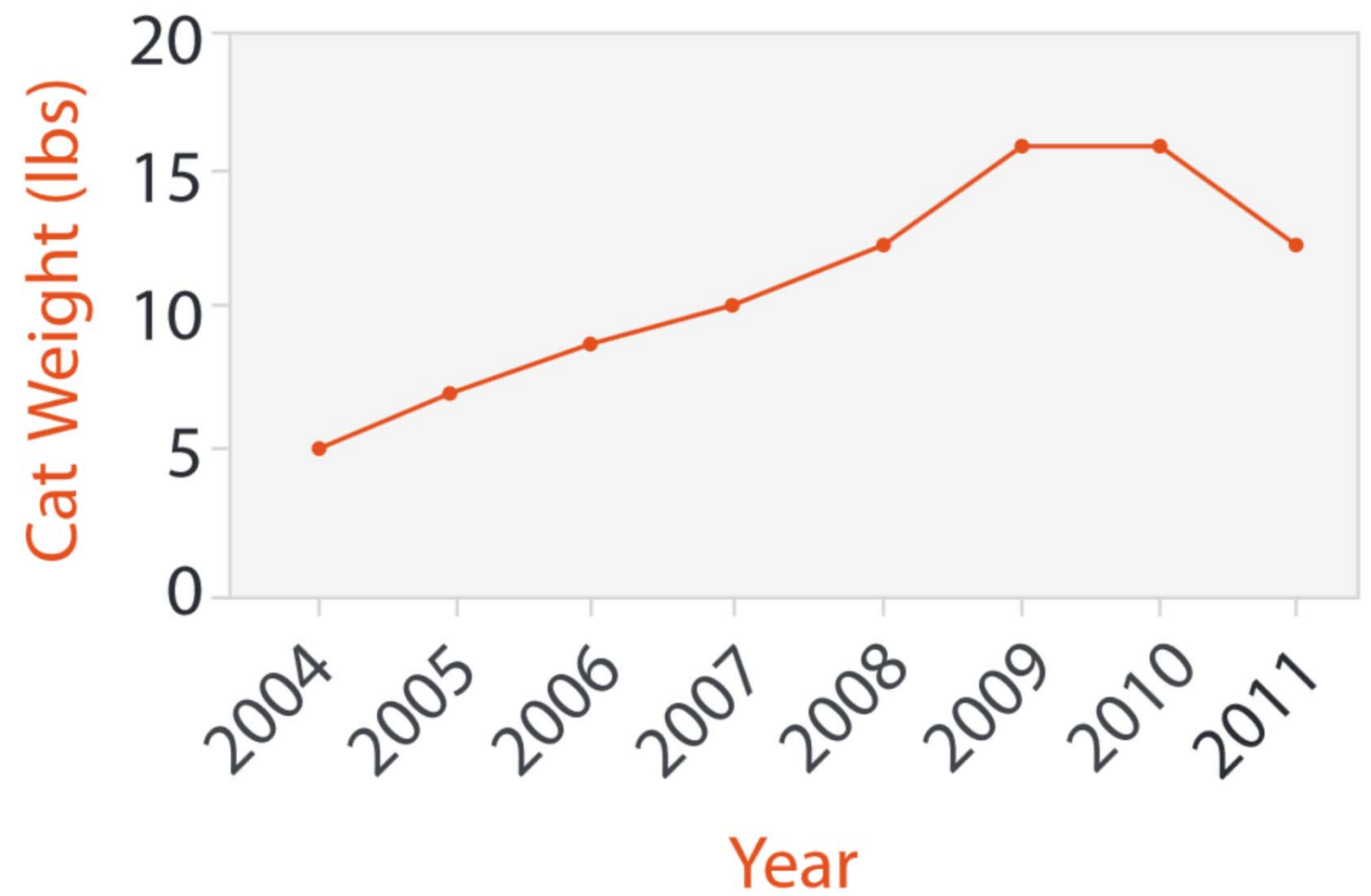
line chart / line graph

- Line emphasizes ordering & continuity
 - Do not use for categorical data!!
 - Also: careful with data gaps



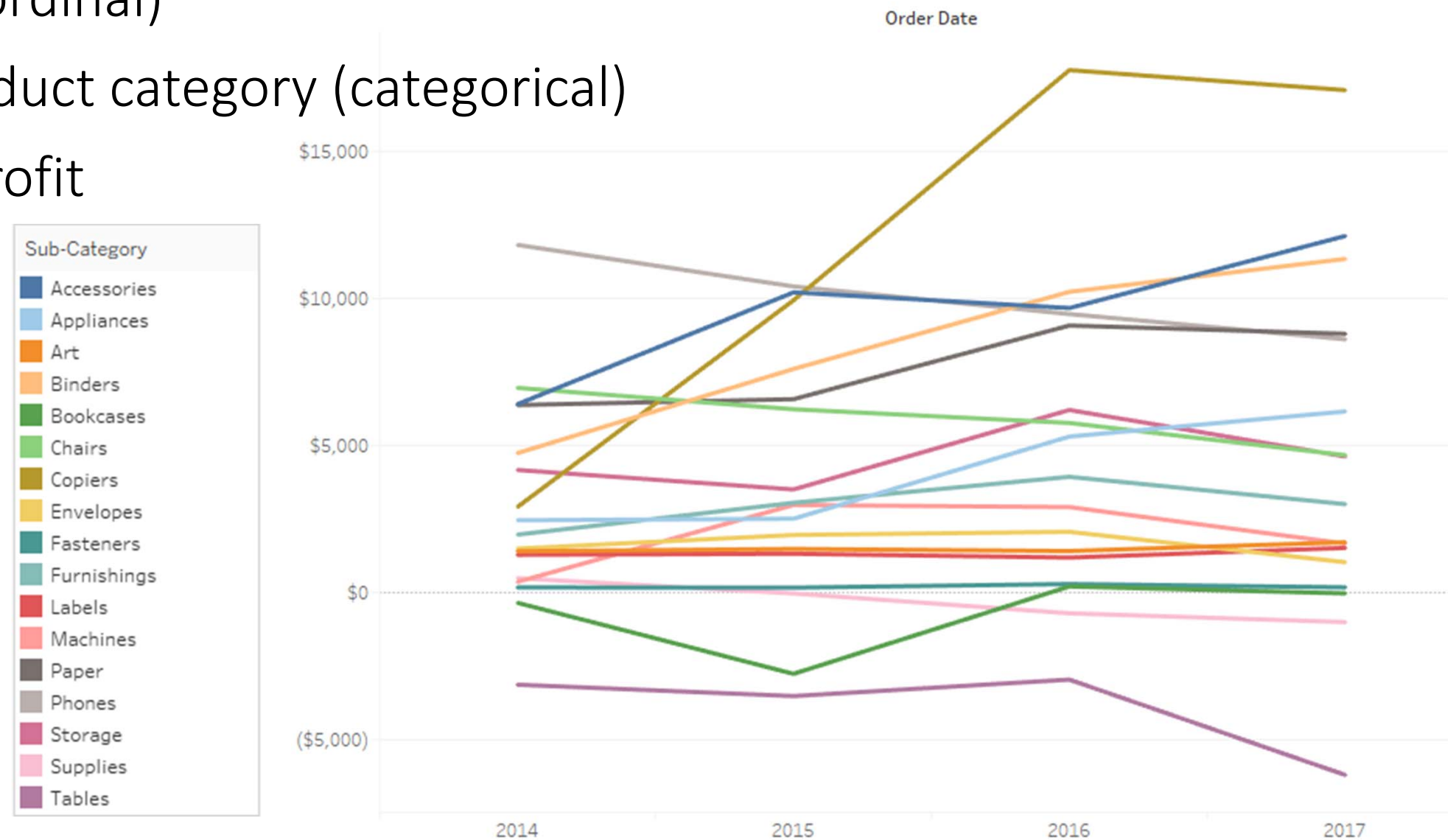
line chart / line graph

scale Primary attribute (ordinal): hundreds of levels

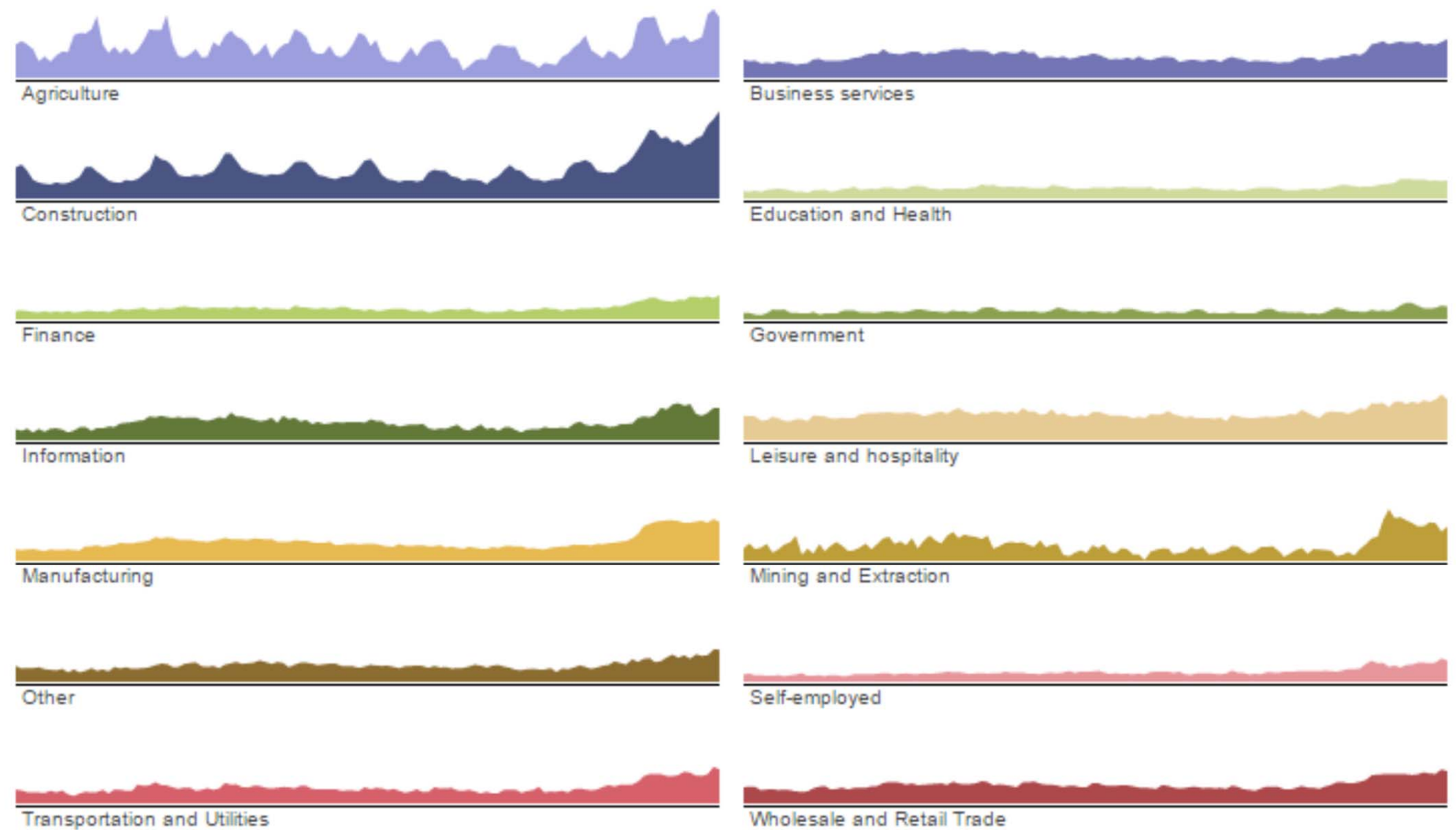


showing multiple attributes through line charts

- Primary attribute: year (ordinal)
- Secondary attribute: product category (categorical)
- Quantitative attribute: profit



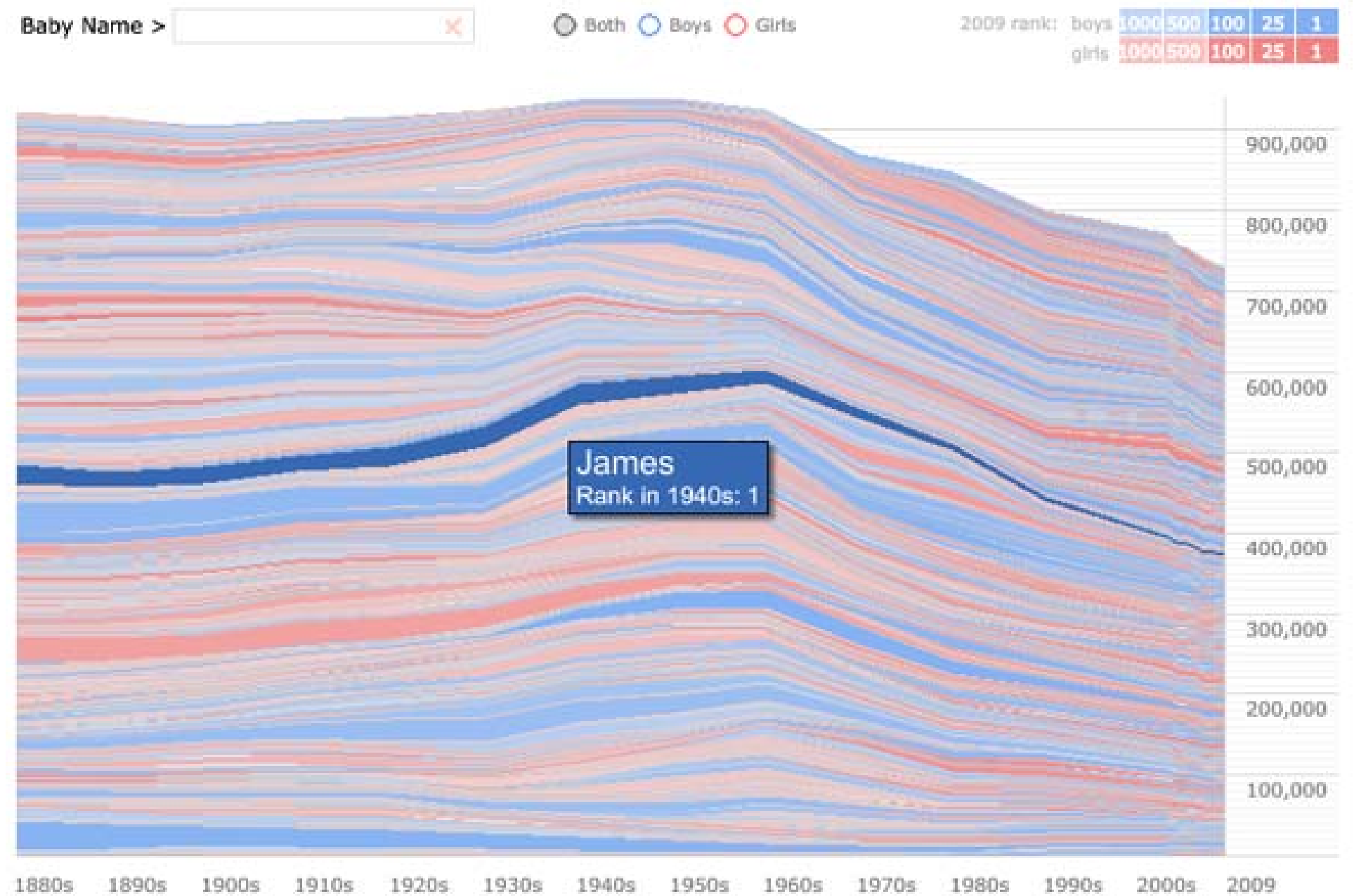
multiple area charts



US unemployment rates 2000 - 2010

<http://homes.cs.washington.edu/~jheer//files/zoo/ex/time/multiples.html>

stacked area graphs



Name Voyager, 2005

<http://www.bewitched.com/namevoyager.html>

next week

- Monday: Tableau Tutorial I
- Thursday: more about basic visualization techniques
- Why? → Activities and goals that can be supported via visualization
 - What tasks can visualizations support?
 - Munzner, Chapter 3